Guidance on Evaluation of Data from the Developmental Neurotoxicity (DNT) In-Vitro Testing Battery

Original Draft: July 2021

Revised Draft: October 2022

Foreword

Developmental neurotoxic chemicals are a diverse set of substances that have the potential to interfere with the normal development of the nervous system, which, if perturbed without compensation, may lead to adverse effects on the normal development of nervous system structures and/or functions. The current OECD test guideline (TG) 426 requires assessing the impact of prenatal and postnatal exposure on physical and neurodevelopmental landmarks, clinical observations, and behavioural and neuropathological endpoints (OECD 2007). Use of the in vivo test guideline has been limited due to extensive time and resource costs, the number of animals used, and lack of chemical alerts that trigger it. This limited testing, coupled with an increasing need to assess the hazards of hundreds of pesticides and thousands of industrial chemicals, has resulted in calls for the development and use of in vitro methods. International efforts have led to the development of the Developmental Neurotoxicity In Vitro Battery (DNT IVB). The DNT IVB covers mechanisms directly involved in DNT as depicted in key neurodevelopmental processes and not mode of actions that may lead to secondary effects on neurodevelopment (e.g., endocrine and immune mode of actions). The purpose of the present document is to provide guidance on the evaluation of data developed with the DNT IVB. The major aims are to describe the assays that comprise the battery in terms of neurodevelopment, provide criteria that allows evaluation of the relevance of the data to developmental neurotoxicity, and to assist in the determination of the degree of certainty in any positive or negative findings to better inform use of DNT in vitro data in regulatory hazard determinations. Although a list of regulatory needs identified by international working groups is provided, it is expected that the different needs of various regulated chemical sectors and regulatory frameworks will dictate their use of the DNT IVB.

The original draft of this document was developed by Kevin M. Crofton and William R. Mundy who served as consultants to the Danish Technical University under contract with EFSA (Contract number: NP/EFSA/PREV/2019/01) and published in 2021 (Crofton and Mundy, 2021). A second version was developed after review and revisions by the co-Chairs of the Expert Group, Andrea Terron (EFSA), Timothy Shafer (US EPA), with input from Expert Group members Susanne Hougaard Bennekou, Stefan Masjosthusmann, Ellen Fritsche, Anna Price, Kate Willet, Francis Bailey, Rex FitzGerald, the Secretariat, and the external experts Jack Fowle and Pamela Lein. The current version is the outcome of two commenting rounds by the members of the OECD Expert Group on the DNT IVB and one commenting round by the Working Party of the National Coordinators of the Test Guidelines Programme (WNT).

- Purpose – The purpose of this document is to provide guidance on how to evaluate data from the DNT IVB (e.g., hit vs non-hit, uncertainties, biological coverage). It is not intended to guide the use of results in human hazard and risk assessments. Specific criteria for such use will likely be available in specific regulations or developed by regulatory authorities who will determine acceptability based on their needs.
- Assay Inclusion: The assays included in the current iteration of the DNT IVB are those that: 1) were deemed ready for use in screening and prioritisation (Fritsche et al. 2017; Bal-Price et al. 2018; Sachana et al. 2019); 2) have been tested using a common set of chemicals (see Appendix E); 3) data have been analysed using the US EPA's ToxCast Pipeline (TCPL) (ToxCastDB v3.5, ToxCast Website); and 4) have detailed descriptions in the ToxTemp (Krebs et al. 2019) format in Appendix B.
- Chemicals: There have been a total of 476 compounds tested in one or more DNT IVB assays, and 81 compounds tested in all the assays of the DNT IVB. The chemicals were chosen based on criteria that included: 1) assay-specific chemicals used in assay development; 2) a list of possible positive and negative developmental neurotoxicants; and 3) chemicals important to funding agencies.
- Negative findings: At this time negative results from the DNT IVB should not be interpreted as a lack of DNT potential. This is due to the uncertainties associated with the in vitro methods used, as well as the lack of coverage for some critical neurodevelopmental processes.
- Adversity: At this time interpretations of adversity is outside the scope of the document. As with many in vitro test methods, there is a lack of empirical data to correlate specific levels of changes in the assays with known alterations in in vivo neurodevelopmental outcomes.
- Validation: A classic OECD 34 validation of the DNT IVB (e.g., inter-laboratory testing of all assays using defined lists of positive and negative compounds) has not been conducted. This uncertainty is clearly stated in the GD and should be considered by regulatory authorities for use in a fit-for-purpose manner driven by their decision needs. Information on the level of validation of the individual assays can be found in the ToxTemp files in Appendix B.

Abbreviations Table	
AChE	Acetylcholinesterase
AO	Adverse Outcome
AOP	Adverse Outcome Pathway
ATP	Adenosine Triphosphate
BMC	Benchmark Concentration
BMR	Benchmark Response
CV	Coefficient of Variation
CYPs	Cytochromes
DNT	Developmental Neurotoxicity
DNT IVB	Developmental Neurotoxicity In Vitro Battery
EC	Effective Concentration
ECHA	European Chemicals Agency
EFSA	European Food Safety Authority
ESCs	Embryonic Stem Cells
EU	European Union
GD	Guideline
HTS	High-Throughput Screening
IATA	Integrated Approaches to Testing and Assessment
iPSC	Induced Pluripotent Stem Cells
IUF	Leibniz Research Institute for Environmental Medicine
IVB	In Vitro Battery
IVIVE	In Vitro to In Vivo Extrapolation
KE	Key Event
KER	Key Event Relationship
LDH	Lactate Dehydrogenase
MEA	Multi-Electrode Assay
MIE	Molecular Initiating Event
NAMs	New Approach Methodologies
NNF	Neuronal Network Formation
NPCs	Neural Progenitor Cells
NSCs	Neural Stem Cells
NT	Neurotoxicity
NTP	National Toxicology Program
OECD	Organisation for Economic Cooperation and Development
SD	Standard Deviation
SOP	Standard Operating Procedure
TG	Test Guideline
ToxPi	Toxicological Prioritisation Index
(Q)SAR	(Quantitative) Structure Activity Relationships
UKON	University of Konstanz
US EPA	US Environmental Protection Agency
WoE	Weight of Evidence

Table of Contents

1. Introduction

1.a. Purpose

1. Developmental neurotoxicants are a diverse set of substances that have the potential to interfere with the normal development of the nervous system, which, if perturbed without compensation, may lead to adverse effects on the normal development of nervous system structures and/or functions (Mileson and Ferenc 2001; USEPA 1998). Development of the nervous system involves a complex interplay between multiple processes that occur both prenatally and postnatally and are developmental stage and location dependent. This has led to the general observation that the developing nervous system is particularly vulnerable to environmental chemicals (Grandjean and Landrigan, 2006; 2014). Due to the vulnerability and potentially high societal costs of adverse impacts on neurodevelopment, the potential health effects following exposure to environmental chemicals led to development of in vivo testing batteries starting in the 1980's, and development of testing guidelines by the USEPA in 1986, with refinement by OECD in 2007.

2. Current EPA and OECD DNT test guidelines require assessing the impact of prenatal and postnatal exposure on the development of physical and developmental landmarks, clinical observations, behavioural and neuropathological endpoints (OECD 2007). However, there has been limited use of these test guidelines, with a total of approximately 165 chemicals assessed to date using either EPA/OECD DNT or TG443 extended one-generation test guidelines (Crofton 2021; Makris et al. 2009; OECD 2008). This is the result of a number of factors, including: the limited regulatory requirement for DNT testing as compared to some other test guidelines, testing is both time- (e.g. 1-2 years) and resource-consuming, limited triggered testing by chemical alerts, and the need to reduce animal use), (Paparella et al. 2020; Tohyama 2016; Tsuji and Crofton 2012). This limited testing, coupled with an increasing need to assess the hazards of hundreds of pesticides and thousands of industrial chemicals, has resulted in calls for the development and use of New Approach Methodologies (NAMs) that reduce the use of animals (Kavlock et al. 2018) and are efficient and predictive for DNT testing (Bal-Price et al. 2015; Barbosa et al. 2015; Crofton et al. 2011; Fritsche et al. 2017; Lein et al. 2005; Fritsche et al., 2017).

3. The output from the 2016 OECD/EFSA DNT workshop was a consensus that an existing in vitro testing battery (IVB) could immediately be used for screening and prioritisation (OECD 2017a), and that further work was needed to gain international acceptance for hazard identification and characterisation. This formed the basis of several global projects. These include a long-running project by the USEPA to develop and validate in vitro DNT methods (USEPA 2020a), a project led by EFSA to generate more data from the DNT IVB (Masjosthusmann et al. 2020), one led by the Danish EPA to test high priority pesticides (https://mst.dk/service/nyheder/nyhedsarkiv/2018/jun/forskningsprojekter2018/), and another by the US National Toxicology Program (Behl et al. 2019) to develop a rapid and cost-effective screening strategy to prioritise replacements for classes of chemicals (e.g. flame retardants). Furthermore, in response to the outcome and the recommendations of this workshop, in 2017 the OECD convened an international Expert Working Group to develop a framework for evaluation of in vitro DNT testing and use of DNT IVB data in Integrated Approaches to Testing and Assessment (IATA) based on case studies using DNT IVB data.

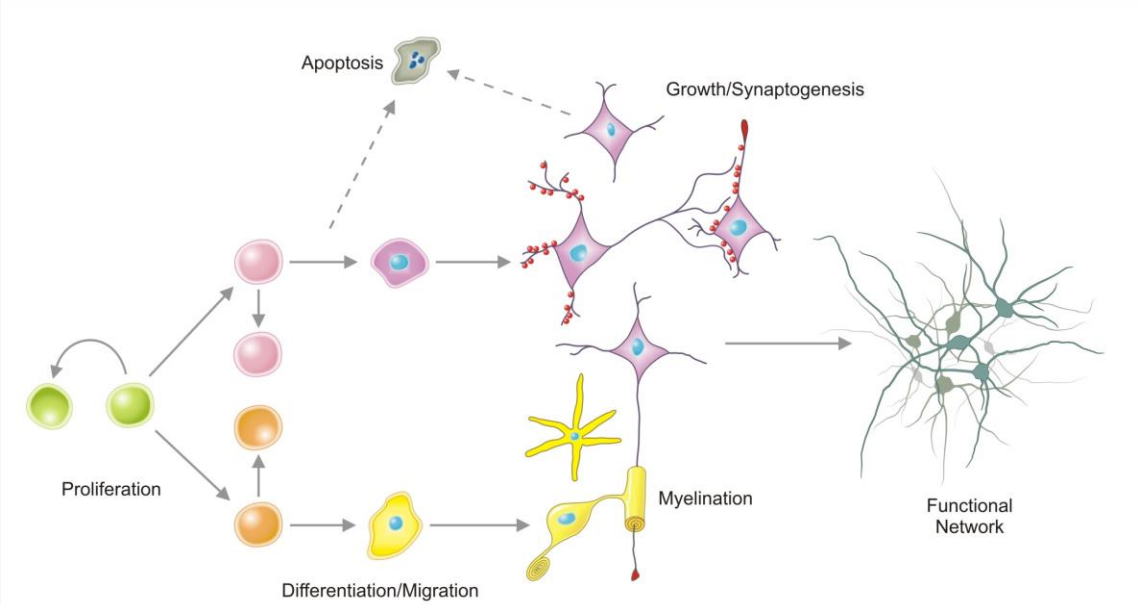
4. The overall purpose of the present document is to provide guidance on the evaluation of data developed with the DNT IVB. The major aims are to describe the assays that comprise the battery in terms of neurodevelopment, provide criteria that allows evaluation of the relevance of the data to developmental neurotoxicity, and to assist in the determination of the degree of certainty in any positive or negative in vitro bioactivity to better inform use of DNT in vitro data in regulatory hazard determinations. The aim of this document is not to guide the use of results in human hazard and risk decisions. Specific criteria for such use will likely be available in specific regulations or developed by regulatory authorities who will determine acceptability based on their needs.

1. b. The Developmental Neurotoxicity In Vitro Battery (DNT IVB)

5. The DNT IVB developed for the EFSA funded research project was based on the OECD/EFSA review of available in vitro DNT assays (OECD, 2017a; Fritsche et al. 2017). This review proposed a set of assays that are not based on molecular targets, as no comprehensive list of targets is currently known, but instead on a number of fundamental neurodevelopmental processes (Figure 1.1, see Section 2.b. for details). This battery of assays has been reviewed, by expert panels (Fritsche et al. 2017; Sachana et al. 2019; Bal-Price et al. 2018), and deemed ready for use in the regulatory arena.

6. The use of the assays in DNT IVB that measure changes in neurodevelopmental processes is based on the assumption that changes in these processes will reflect the integration of chemical disruptions in multiple up-stream molecular events (Lein et al. 2007; Lein et al. 2005; Radio and Mundy 2008). This battery contrasts with

Figure 1.1. Fundamental neurodevelopmental processes necessary for proper nervous system development. In vivo studies have shown that several biological processes at the cellular level are essential for nervous system development. Neural stem cells (NSC, green) proliferate and differentiate into multiple types of neural progenitor cells (NPCs) including neuronal progenitors (light purple) and glial progenitors (orange). These proliferate, migrate, and differentiate into neurons (purple) and glia (yellow). As cells mature, they extend neurites and form synapses (red). Surplus cells undergo apoptosis (grey). When these events happen in a coordinated fashion, cell-cell interactions result in a functional neuronal network (olive). (From Aschner et al., 2017, courtesy of William Mundy and John Havel).



many other in vitro batteries that cover a single cellular pathway (e.g., oestrogen receptor, Judson et al., 2015), in that it seeks to predict the impact of xenobiotic exposures on the development of an entire organ (i.e., brain). To date about 476 chemicals have been tested in up to 17 assays in the DNT IVB (see Appendix E), using animal- and human-based cell cultures that can measure changes in proliferation, differentiation, apoptosis, migration, neurite formation, synaptogenesis, and neural network formation (Masjothusmann et al. 2020; Harrill et al. 2018; Shafer et al., 2019). Each assay includes assessment of cell viability and/or cytotoxicity. This battery was developed based on current knowledge; it should be considered an "evergreen" document with improvements and revisions occurring as new assay methods (e.g., Culbreth et al., 2022; Loser et al 2021) and more chemical test data become

available to improve predictive capacity of the whole test battery, and as the reliability of each individual test method is improved through further validation work.

1.c. Target Uses

7. It is important to note that currently there is not sufficient evidence that the DNT IVB can replace the use of OECD TG426 and the DNT Cohorts 2A and 2B in OECD TG443 for derivation of all hazard-based decisions. Instead, targeted DNT IVB testing should be guided by a problem formulation approach based on regulatory needs using the IATA framework (OECD 2016; Sachana and Leinala 2017; Sakuratani et al. 2018). Below is a list of needs for some regulatory frameworks identified by international working groups (Fritsche et al., 2017; Bal-Price et al., 2018; Sachana et al., 2019).

- Follow-up testing for positives identified by quantitative structure activity relationships ((Q)SAR), read-across and other predictive computational models of developmental neurotoxicity
- Screening for prioritisation
    - Screening of large numbers of chemicals that lack or have limited data on DNT (e.g., Tox21/ToxCast, Health Canada prioritisation scheme).
    - Screening of small numbers of structure/class specific chemicals (e.g., Behl et al. 2019).
- Single chemical hazard assessments
    - When no in vivo DNT data exists, in some regulatory frameworks data from the DNT IVB may be used to determine if, and what, follow-up testing (e.g., orthogonal assays, alternative species (e.g., zebrafish), or a guideline study) could be conducted.
    - If existing in vivo DNT data is equivocal, data from DNT IVB testing could be used to inform the Weight of Evidence (WoE)-based IATA assessment for DNT.
    - When DNT in vivo data exists and is negative, but concern exists from new findings or novel in vitro assays (i.e., not currently covered in the DNT IVB assay), the DNT IVB may be used to inform the WoE .
    - When data from an MIE[1] based assay or alternative species assay (e.g., zebrafish) data exist, a regulatory choice could be to run the DNT IVB battery to inform the WoE-based assessment.

1.d. Target Chemicals

8. The target chemicals for possible use in the DNT IVB are industrial chemicals, pesticides, and environmental contaminants, including metabolites and environmental degradates. There are known limitations of chemical testing capability in in vitro assays related to the physiochemical characteristics (e.g., volatility, high reactivity, hazard level, and limited solubility in assay appropriate media) of the test compounds (Richard et al. 2020; Richard et al. 2016). These types of chemical limitations for the individual assays in the DNT IVB should be described in assay documentation as per OECD GD211 (OECD 2017b) and/or the ToxTemp format (Krebs et al. 2019). Descriptions for assays in the DNT IVB, using the ToxTemp format, can be found in Appendix B.

9. Currently, a total of 81 compounds have been tested in all 17 assays in the DNT IVB.[2] A total of 97 compounds have been tested in 14 or more assays, and 331 compounds have been tested in at least four assays. A total of approximately 476 compounds have been tested in one or more of the assays in the DNT IVB. Appendix E provides a list of which assays have been conducted for each chemical.

[1] More information on AOP Framework can be found in Section 2 and Appendix C.
[2] As of February 2022.

10. ***1.e. Aims and Context***The purpose of this document is to provide guidance on how to evaluate data from the DNT IVB (e.g., hit vs non-hit, uncertainties, biological coverage). It is not intended to guide the use of results in human hazard and risk assessments. Specific criteria for such use will likely be available in specific regulations or developed by regulatory authorities who will determine acceptability based on their needs. Nor is it intended to provide guidance for all in vitro assays purported to detect DNT (e.g., available in published literature). The structure of this guidance should be expanded in the future to encompass improvements to the current assays in the DNT IVB, updated validation information, and/or new and novel assays that complement or expand the DNT IVB as it currently exists.

11. The assays in the DNT IVB were chosen based on international discussions on the ability of existing in vitro DNT assays to measure neurodevelopmental processes and the readiness of the assays, at a screening level, to test large numbers of chemicals (Fritsche et al. 2017; Bal-Price et al. 2018; Sachana et al. 2019). The specific assays included in the current iteration of the DNT IVB are those that: 1) were deemed ready for use in screening and prioritisation (Fritsche et al. 2017; Bal-Price et al. 2018; Sachana et al. 2019); 2) have been tested using a common set of chemicals (see Appendix E); 3) data have been analysed using the US EPA's ToxCast Pipeline (TCPL), a common normalization and dose-response modelling approach (ToxCastDB v3.5, ToxCast Website); and 4) have detailed descriptions in the ToxTemp format in Appendix B.

12. This guidance was developed based on the testing of a set of chemicals that is limited compared to the overall chemical universe (Richard et al. 2020), and a set of assays that do not include all critical processes within the developing nervous system (see Section 2.b). Thus, there are uncertainties in the predictive accuracy of the DNT IVB that cannot currently be quantified. This uncertainty will only be reduced with the testing of additional chemical libraries and by integrating additional assays in the DNT IVB (see Section 4.e). This uncertainty needs to be considered when using data from the DNT IVB within any regulatory assessment (see Section 4.f).

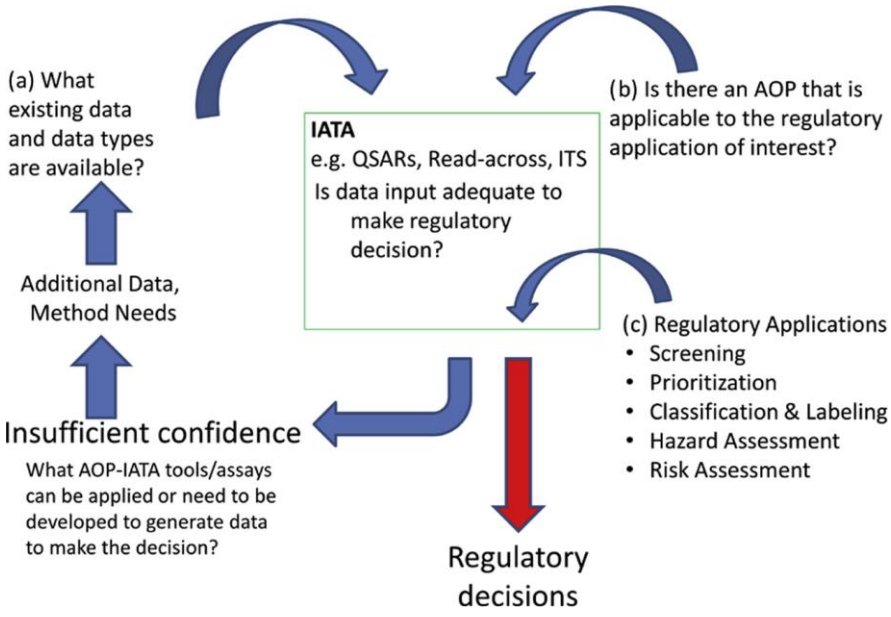
13. The remainder of this document contains a background on DNT assays, a brief description of the assays in the battery, elements for evaluation and use in WoE considerations for both individual assays and the battery, and a brief review of available case studies. This document is intended to briefly review the science that supports the DNT IVB and to provide guidance on how to evaluate data from the DNT IVB.

2. Background

2. a. Integrated Approaches to Testing and Assessment (IATA) for DNT

14. IATA is a framework developed by OECD that allows for the integration of all available hazard and exposure data, including in silico, in chemico, in vitro and in vivo, for use in chemical regulatory assessments (OECD 2016). IATAs are "pragmatic, science-based approaches for chemical hazard characterisation that rely on an integrated analysis of existing information, with optional use of the adverse outcome pathway (AOP) framework, coupled with the generation of new information if necessary. IATAs follow an iterative approach to answer a defined question in a specific regulatory context, taking into account the acceptable level of uncertainty associated with the decision context." (OECD 2016; Sachana and Leinala 2017; Sakuratani et al. 2018).

Figure 2.1. A conceptual framework for an AOP-informed IATA to support regulatory assessments (from Tollefsen et al. 2014).



15. Problem formulation is the first step in IATA development and is critical to the use of the DNT IVB. This involves an understanding of the scope and needs of the risk assessment objective, data requirements, and the level of acceptable uncertainty associated with the decision being made (Tollefsen et al. 2014).

16. Regulatory needs in an IATA can range from screening of large chemical classes for prioritisation of follow-up testing, to regulatory acceptance of the use of new chemicals (see also Section 1.d.). An IATA is based on an evaluation of all existing evidence and may include use of AOPs or other frameworks. Figure 2.1 illustrates these approaches, which can be iterative if necessary. If the uncertainties in the IATA are acceptable within the context of a given regulatory need, the IATA may be used for that regulatory purpose. If the uncertainties in the IATA are not acceptable due to insufficient confidence of the totality of available information, additional data that can address the uncertainties should be identified and input into the IATA. One basic approach of IATAs is to apply the AOP framework to integrate and organise the various data sources. AOPs allow the mapping of data to molecular initiating events (MIEs) and key events (KEs) at the molecular, cellular, organ levels, that lead to an adverse outcomes (AOs) at organismal and population levels (Ankley et al. 2010; OECD 2016). A diagram that illustrates potential components of a DNT AOP is shown in Figure 2.2.

17. Within IATA it is important to clearly articulate the uncertainties in the use of the AOP framework for DNT. First, only a limited number of putative and reviewed AOPs exist for DNT (Bal-Price et al. 2015; Spinu et al. 2019) (see Appendix C for an updated list). Second, most of the assays used in the DNT IVB do not measure MIEs, but instead downstream KEs. While this does not impact the use of DNT IVB data in AOP-based IATAs, it does limit use of DNT IVB data in development of in silico models that are based on chemical interactions with the MIE (i.e., initial biological target) as opposed to downstream KEs. Third, the basic assumption underlying the DNT IVB is that if a chemical disrupts neurodevelopmental processes in vitro, then it has the potential to do the same in vivo (Lein et al. 2007; Lein et al. 2005; Radio and Mundy 2008). Understanding these intrinsic uncertainties is critical to inform regulatory authorities in the use of data derived from the DNT IVB. As more data are generated using the DNT IVB, and with further AOP development and improved understanding of mechanisms underlying developmental neurotoxicity (see Section 4), the uncertainties in the use of DNT IVB will be reduced.

18. A description of existing DNT AOPs and the relationship to KEs measured in the DNT IVB assays is found in Appendix C.

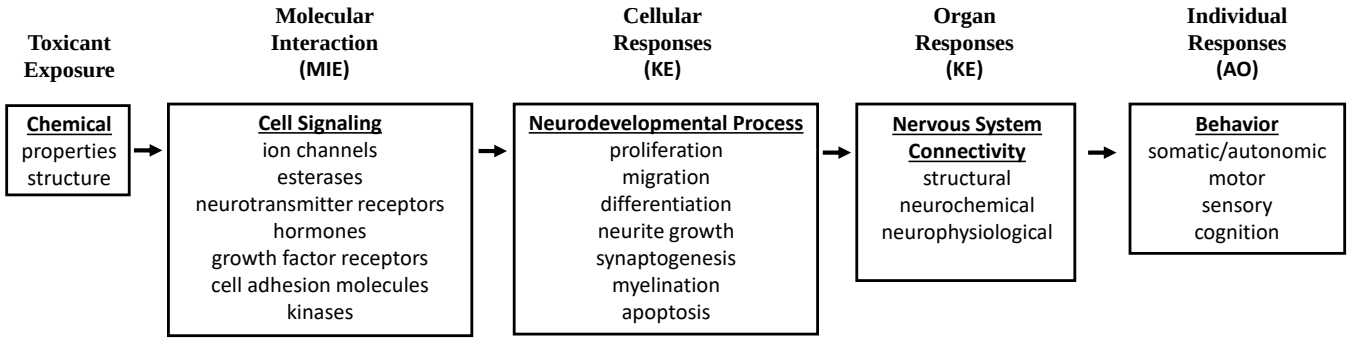


Figure 2.2. Diagram illustrating potential components of an Adverse Outcome Pathway (AOP) for Developmental neurotoxicity. Pathway showing the relationship between molecular initiating events (MIEs) such as altered cell signaling, and key events (KEs) at the cellular level (neurodevelopmental processes) and organ level (nervous system connectivity), which may lead to an adverse outcome (AO) for the individual.

2. b. Developmental Neurotoxicity In Vitro Battery (DNT IVB)

In vivo development of the nervous system proceeds through a series of coordinated biological processes that are essential for the formation of normal brain structure and function. While the timing, duration, and spatial location of these processes differ between species, the fundamental processes are remarkably conserved across invertebrate, non-mammalian vertebrate and mammalian organisms (Sanes et al. 2011). These fundamental neurodevelopmental processes include progenitor cell proliferation, differentiation into neuronal and glial cells, migration, apoptosis, axonal and dendritic outgrowth, myelination, synapse formation and formation of functional neural networks (Figure 1.1). In vivo, perturbation of one or more of these neurodevelopmental processes during nervous system development may, if not compensated, result in adverse effects on brain structure and/or function (Rice and Barone 2000; Rodier 1994; De Vries, 2004). ***Thus, use of the DNT IVB is based on the assumption that chemicals affecting one or more of these neurodevelopmental processes in vitro have the potential to do so in vivo (Lein et al. 2007; Lein et al. 2005; Radio and Mundy 2008).*** This does not, in any way, imply that in vitro changes in these processes, de facto, predict an adverse outcome in vivo. Some aspects of these neurodevelopmental processes can be examined as discrete, measurable events in vitro that provide biologically relevant information about chemical perturbations at the cellular level. In terms of the AOP framework, the assays which comprise the DNT IVB can be considered as cellular/organ level KEs in an AOP for DNT (Figure 2.2). An advantage of measuring cellular-level neurobiological processes (e.g., migration) in cell-based assays (i.e., assays in which cell cultures are exposed to the test chemical) is that the intact cell integrates potential chemical actions at multiple upstream targets (MIEs) and cellular pathways (KEs) (Lein et al. 2007; Lein et al. 2005; Radio and Mundy 2008). For use in hazard identification and characterisation and chemical

screening, cell-based assays “cast a wide net” (Cooper et al. 2017), in that they can detect chemical effects in the absence of prior knowledge of MIEs. In addition, use of the AOP framework for selection, and development of assays in the DNT IVB identifies the potential links between MIEs, KEs and AOs and provides a mechanistic basis supporting the relevance of the data derived from the DNT IVB. Concurrent measurement of cell viability and/or cytotoxicity in cell-based assays allows for the evaluation of data in terms of a “selective effect” of a chemical on a neurodevelopmental process (i.e., changes occurring in the absence of an overall effect on cell health or viability), further supporting the relevance of in vitro data to perturbation of neurodevelopment.

19. Due to the many known and unknown modes of action for chemical effects on nervous system development, any one individual assay is unlikely to detect all developmental neurotoxicants. Rather, a battery of assays covering multiple neurodevelopmental KEs has been proposed that should increase the ability to detect developmental neurotoxicants. Numerous test methods (assays) have been developed for quantitative assessment of chemical effects on aspects of neurodevelopmental processes in vitro (e.g., Fritsche et al. 2017; Schmidt et al. 2017). These assays use a variety of neural cell types derived from human and animal sources (test systems) and measure molecular, morphological, and electrophysiological endpoints. In most cases, methods that are amenable to medium- or high-throughput testing are used. Details on published methods for measuring the neurodevelopmental processes at the cellular level and their application for use in regulatory DNT testing have

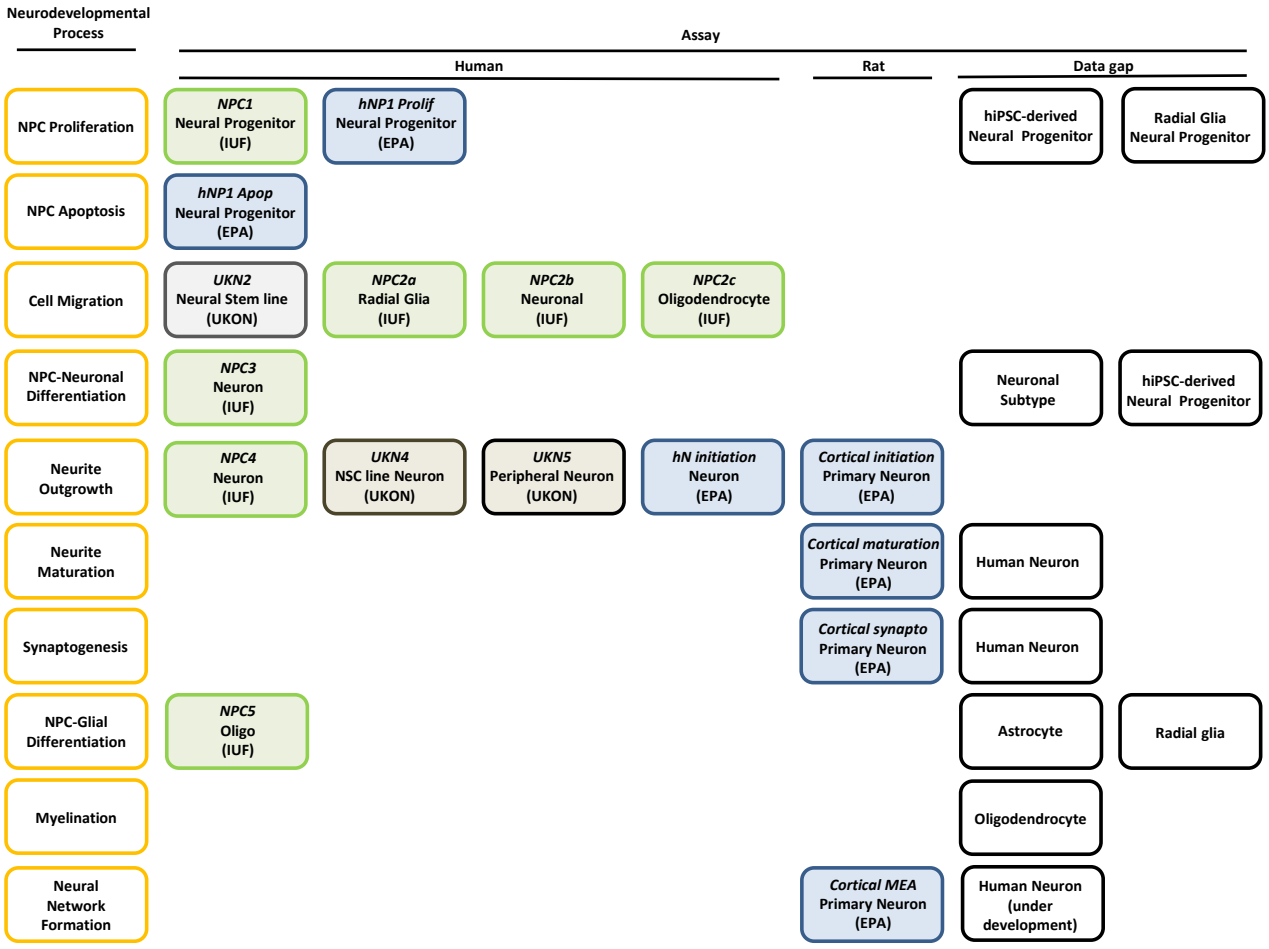


Figure 2.3. Assays in the current DNT IVB and assays identified as high priority for development (modified and adapted from Masjosthusmann et al., 2020). Assays are grouped according to the neurodevelopmental process evaluated (rows) and test system used (columns). Each assay box lists the abbreviated assay name, cell type and assay development laboratory. Assays which need further development, and not included in the current DNT IVB, are identified as data gaps. Each assay is represented as a box which lists the test method name (italics), the test system (cell type used), and the home institution of the developer (IUF =Leibniz Research Institute for Environmental medicine; UKON =University of Konstanz; EPA= US Environmental Protection Agency.

been reviewed (OECD, 2017a). The conclusion of that review was that for many of the neurodevelopmental processes, in vitro assays are available that permit quantitative assessment of relevant endpoints (i.e., measured variable). Due to advances in stem cell biology that have made human cells more accessible, test systems using human-derived neural cells are becoming more common; such test systems are preferred since they should minimise interspecies extrapolation issues. However, human neural cells can be difficult to culture, and the developmental timeline of such cells may be much longer than rodent-derived neural cells (Barry et al. 2017; Nimtz et al. 2020; Odawara et al. 2016). For some processes (e.g., synaptogenesis), assays incorporating human-derived neurons and glia are being developed (Nimtz et al. 2020), but currently only assays using rodent-derived cell cultures were deemed ready for screening. An international expert group on DNT completed an analysis that scored and ranked the available in vitro assays for their readiness for use in chemical screening/prioritisation or risk assessment (Bal-Price et al. 2018). Readiness criteria included assay descriptions (e.g., purpose, relevance, cell

type, reference chemicals, exposure, technical limitations), performance (e.g., repeatability, variability, sensitivity/specificity), data evaluation (e.g., dynamic range, curve fitting, benchmark response) and application (e.g., chemicals, relevant pathways/AOPs, prediction models). Additional information used to evaluate assay readiness included results from assays that have already screened tens to hundreds of chemicals. Based on the recent reviews of DNT in vitro assays (Bal-Price et al. 2018; Fritsche et al. 2017), a battery of assays (DNT IVB) has been assembled that covers a majority of the key neurodevelopmental processes (Masjosthusmann et al. 2020). Figure 2.3 illustrates the assays in the current battery (in total 17 assays) organised by key neurodevelopmental processes (rows). The figure indicates the species origin of the test system used and lists neurodevelopmental processes where assays should be a high priority for development. In many cases, multiple test methods are available to assess a single neurodevelopmental process. These assays are used currently used in the battery in a complementary manner (as opposed to orthogonal assays) (see Textbox 1).

20. However, due to differences between test systems (e.g., neural stem cell versus primary neural cell cultures) and test methods (e.g., developmental timing, assay duration, endpoints, exposures), different outcomes from assays for the same neurodevelopmental process are possible. Three assays in the DNT IVB that measure migration are a case in point. The UKN2 assay utilises neural crest cells, a type of neural stem cell that migrates towards tissues outside the nervous system during early embryogenesis and can differentiate into both neural and non-neural cells (Mayor and Theveneau 2013). In contrast, the NPC2a assay uses radial glia cells that normally migrate as a prerequisite for cortex formation (Borrell and Gotz 2014), and the NPC2c assay uses oligodendrocytes, a more mature cell that arises and migrates later in development (Barateiro and Fernandes 2014). The biology that regulates migration is different in these test systems (Minoux and Rijli 2010; Ortega et al. 2012; Sild and Ruthazer 2011), providing different targets and pathways for testing chemical interactions. Another example is the neuronal network formation assay (NNF). The cortical cell-based Multi-Electrode Assay (MEA), and Human NNF assay being developed, use the same basic cell types (neurons and astrocytes), but from different

Textbox 1.

Complementary Assays: Assays that measure similar endpoints and are conducted in conjunction rather than as follow-ups to primary screening. These assays may use similar or different technologies or test species and may evaluate different levels of biological complexity. Such assays can provide additional certainty when they are all positive or negative. However, different findings between the assays are not necessarily indicative of a false positive in that said differences could be due to differences in assay methods such as cell type, species, level of biology evaluated, or assay technology.

Orthogonal Assays: An assay performed following the primary assay to differentiate between compounds that generate false positives from those substances that are genuinely active against the target. Such assays use different reporter or assay technologies in an effort to confirm that activity of the compound is directed toward the biological target of interest, or if negative suggest that the original substance activity was most likely assay format-dependent and not specific to the biology of interest. In some cases, orthogonal assays can be run at the same time (adapted from Thorne et al., 2010).

species (rat primary neocortex versus human induced pluripotent stem cell- and primary astrocyte-derived). Test results from these assays may be different due to differences in developmental timing (Saavedra et al. 2021) as well as differences between neural cells derived from different brain regions and in the pathways and proteins expressed in rats versus humans (Robbins et al. 2010). Thus, to facilitate evaluation of the results in terms of the stated test purpose it is important that both the test system and test method be clearly described and annotated by the assay developer. This information is documented in an organised format developed for cell-based assays (see below) and is found Appendix B. Table 2.1 provides a general description of the types of cell cultures used as test systems in the current DNT IVB. Each of these different types of neural cultures can be derived from rodent or human tissue.

Table 2.1. General descriptions of neural cell cultures used as test systems in the current DNT IVB	
Type of Cell Cultures	**Description**
Neural Stem Cells (NSC)	Cells that can self-renew and proliferate indefinitely. They produce progeny cells (e.g. neural progenitor cells such as radial glia) that ultimately give rise to a majority of cells in the nervous system. They can be derived from embryonic stem cells (ESC) or induced pluripotent stem cells (iPSC).
Neural Progenitor Cells (NPC)	Derived as progeny of neural stem cells or directly from fetal brain tissue, they have a more limited proliferative ability. They can terminally differentiate into neurons (including many subtypes) and glia (astrocytes and oligodendrocytes).
Neurospheres	3D culture of neural stem cells and neural progenitor cells that grow in small clusters and can terminally differentiate into a mixed culture of neurons and glia.
Primary Cells	Immature neurons and glia derived directly from brain tissue.

21. Test methods for the DNT IVB are designed to model aspects of one or more neurodevelopmental processes in vitro. The primary purpose of each method is to evaluate whether test chemicals can alter these processes as a function of development in vitro. A general description of each neurodevelopmental process modelled in the DNT IVB along with potential consequences of chemical disruption and endpoints measured is listed in Table 2.2. Note that these are neurodevelopmental processes that may not necessarily reflect developmental timing, location, or neuronal subtype aspects.

22. In vitro DNT test methods have been reviewed (Bal-Price et al. 2018; OECD 2017a), and those that constitute the current DNT IVB are presented in Table 2.3, which is organised by the neurodevelopmental process being modelled

Table 2.2. General descriptions neurodevelopmental process modelled in the DNT IVB	
Process	**Description**
Proliferation	Division of neural stem cells and neural progenitor cells resulting in an increase in cell number. Changes in proliferation can result in an incorrect cell number (increase or decrease) and altered brain growth. It is measured directly by assessing the number of cells undergoing DNA replication or inferred by measuring a change in cell number over time.
Apoptosis	Programmed death of cells resulting in a decrease in cell number. Changes in apoptosis can result in an increase or decrease in cell number and altered brain growth. It is measured by assessment of cell nucleus morphology or detection of biochemical markers specific to the apoptosis pathway.
Migration	Movement of neural progenitor, glial or neuronal cells from their point of origin to a final position. Changes in migration can result in cells in the wrong position and abnormal brain structure. It is measured by assessing the number of cells moving into a defined area, or the distance moved by individual cells.
Neuronal Differentiation	Process in which a neural progenitor cell changes to a specific type of neuron. Changes in differentiation can result in altered cell numbers for specific populations of neurons, changing brain structure and function. It is measured by assessment of the number of cells expressing markers specific for neurons and neuronal subtypes.
Neurite Growth	Outgrowth of morphological processes relatively early in neuronal differentiation. Neurites eventually develop into dendrites or axons. Changes in neurite growth can alter the number and length of axons and dendrites, changing brain structure and connectivity between neurons. It is measured by counting the number of cells elaborating processes or morphological assessment of neurite length.

Neurite Maturation & Synaptogenesis	Maturation of neurites into the specialized processes of dendrites and axons which then form synapses responsible for communication between neurons. Changes in neuronal maturation and synaptogenesis alter neuronal connectivity, changing network formation and brain function. They are measured by morphological assessment of axon and dendrite length and counting of the number of synapses
Glial Differentiation & Maturation	Process in which a neural progenitor cell changes to a specific type of glia (radial glia that support formation of cortical architecture, astrocytes that support neuronal function, and oligodendrocytes that myelinate axons). Changes in differentiation can result in altered numbers of glia and reduced myelination, changing brain structure and function. Differentiation and maturation are measured by assessment of the number of cells expressing markers specific for glial subtypes.
Network Formation	Process in which neurons and glia grow and make functional contacts with each other, exemplified by spontaneous generation and propagation of electrical action potentials within a network. Changes in network formation and function can result in altered neural connectivity and altered brain function. It is measured by electrophysiological assessment of coordinated electrical activity of neurons and glia grown on electrode arrays.
Viability/Cytotoxicity	Test methods for each neurodevelopmental process should also include a concurrent measure of cell viability (or its converse, cytotoxicity) as a baseline for comparison of potential non-specific effects of chemical exposure. Cell viability can be assessed by counting cells with normal morphology (based on cell body and nucleus size), delineation of live/dead cells based on uptake or exclusion of vital dyes, or biochemical assessment of active cell metabolism. Cytotoxicity is typically assessed by measuring parameters associated with loss of cell membrane integrity, including leakage of intracellular proteins and enzymes, or exposure of DNA.

in vitro for chemical testing (i.e., the test purpose). For each test method, a brief description is provided of the test system, exposure scheme, and endpoints measured. To facilitate the evaluation of data from cell-based test methods and potential use in regulatory settings, the OECD has formulated guidance for a more detailed documentation for non-guideline and new approach methods (OECD 2017b). This guidance outlines the standard information that should be provided by developers to assess the quality of data produced and the potential utility in regulatory applications. A more recent and expanded template called ToxTemp, expanding upon OECD GD 211, was developed to guide the user in the details required, specifically for the description of cell-based test systems (Krebs et al. 2019). For the DNT IVB, assay descriptions based on these formats are provided in Appendix B.

Table 2.3. Assays Currently in the DNT In Vitro Battery

Test Method (Assay)	Test System (Cell culture)	Assay Duration/ Chem exposure	DNT Endpoint	Viability/Cytotoxicity Endpoint
Proliferation				
NPC1	human NPC grown as proliferating 3D neurospheres	72 h / 72 h	neurosphere area, BrdU incorporation in dividing cells	Resazurin reduction /LDH release
hNP1 Prolif	human NPC	24 h / 24 h	BrdU incorporation in dividing cells	ATP level
Apoptosis				
hNP1 Apop	human NPC	24 h / 24 h	apoptosis pathway (Caspase) activation	ATP level
Migration				
UKN2	human NSC-derived neural crest cells	72 h / 24 h	number of cells moving into defined area	Calcein-AM vital dye
NPC2a	human NPC grown as differentiated 3D neurospheres	72 h / 72 h 120 h / 120 h	mean distance of radial glia (nuclei negative for neuronal and oligodendrocyte markers) from edge of sphere	Resazurin reduction/LDH release
NPC2b	human NPC grown as differentiated 3D neurospheres	120 h / 120 h	mean distance of tubulin-positive neurons from edge of sphere	Resazurin reduction/LDH release
NPC2c	human NPC grown as differentiated 3D neurospheres	120 h / 120 h	mean distance of O4-positive oligodendrocytes from edge of sphere	Resazurin reduction/LDH release
Neuronal Differentiation				
NPC3	human NPC grown as differentiated 3D neurospheres	120 h / 120 h	number of tubulin-positive neurons	Resazurin reduction/LDH release
Neurite outgrowth				
NPC4	human NPC grown as differentiated 3D neurospheres	120 h / 120 h	neurite length & area	Resazurin reduction /LDH release
UKN4	human NSC-line (v-myc transformed)	72 h / 24 h	neurite area	Calcein-AM vital dye
UKN5	human iPSC-derived peripheral (sensory) neurons	24 h / 24 h	neurite area	Calcein-AM vital dye
hN initiation	human NPC-derived neurons	48 h / 48 h	neurite length	cell morphology
Cortical initiation	rat primary neocortex	48 h / 48 h	neurite length	cell morphology
Neurite Maturation and Synaptogenesis				
Cortical maturation	rat primary neocortex	288 h / 120 h	dendrite length	cell morphology

Cortical synapto	rat primary neocortex	288 h / 120 h	synapse number	cell morphology
Glial Differentiation				
NPC5	human NPC grown as differentiated 3D neurospheres	120 h / 120 h	number of O4-positive oligodendrocytes	Resazurin reduction /LDH release
Neural Network Formation				
Cortical MEA	rat primary neocortex	288 h / 288 h	Action potential spike and burst parameters related to network connectivity	Resazurin reduction/total LDH

388

389 ***2c. Developmental Neurotoxicity (DNT) versus Neurotoxicity (NT) In Vitro***

390 23. DNT testing is based on the observation that the developing brain may be more vulnerable to chemical
391 toxicity than the adult brain. A major reason for increased vulnerability is the dynamic nature of development
392 itself. During this period there is a sequential and coordinated expression of molecular signaling pathways
393 underlying neurodevelopmental processes, and chemical disruption during critical time windows can lead to
394 permanent effects (Rice and Barone 2000; Rodier 1994).

395 24. In vitro assays that measure the impact of chemicals on the nervous system can be divided into two
396 categories: 1) assays for DNT that use test systems that model neurodevelopmental processes, and 2) assays for
397 neurotoxicity (NT) that use test systems that better model the steady state observed in the adult brain and may
398 not model developmental processes. Assays in the DNT IVB are included in the first category and assess the impact
399 of exposures on critical neurodevelopmental processes. For developmental neurotoxicity, in vitro test systems are
400 by definition in a dynamic state, undergoing changes in cellular status over time. For example, in proliferation
401 assays the number of neural progenitor cells increases over time (Moors et al. 2009), while for neurite outgrowth
402 assays neuron cell number is relatively constant but the length of neurite processes is increasing (Harrill et al.
403 2010). This contrasts with the second category, NT assays, which use in vitro test systems that are designed to
404 assess neurotoxicity in cell cultures using "matured" cells that have reached a relatively stable equilibrium. These
405 mature neural cell cultures do not necessarily exhibit the complex and dynamic changes in cell status associated
406 with active neurodevelopment.

407

408 ***2d. Developmental Exposure***

409 25. Proper evaluation of data from DNT in vitro assays requires a good understanding of the exposure regimen
410 used. While there are several factors to consider, the most important is the timing and duration of exposure
411 relative to the timing of dependent variable measurement. A rapid and acute chemical application applied to the
412 in vitro test system over a relatively short time (e.g., seconds to minutes) prior to assessment of the endpoint is
413 not a developmental exposure. Since it does not encompass a period that includes a significant change in the
414 baseline of the endpoint, it is unlikely to measure a change in a neurodevelopmental process, although it may
415 provide information on direct effects of a chemical on cell function.

416 26. *Developmental exposures* require time periods between chemical application and endpoint assessment
417 that coincide with the dynamic changes in cell status associated with neurodevelopmental processes. Since
418 exposures may be different for each test method and test system, they must be documented with reference to
419 the dynamic change in cell status as described above. The actual exposure scheme can range from a single

application without media change at the start of the test to multiple applications (with or without media changes) at regular periods throughout the test. For the DNT IVB, the test duration can range from 24 hours to weeks. For example, proliferation is a process that can occur relatively rapidly in neural progenitor cells. In the DNT IVB, the hNP1 proliferation assay uses a single chemical application at the start of the 24-hr exposure period, during which the cells are undergoing cell division. In contrast, neural network formation is a more prolonged process that includes axon and dendrite outgrowth and synapse formation occurring over days to weeks. In the rat Cortical MEA assay, multiple applications of the chemical are made along with media changes at regular intervals throughout a 12-day test period (Brown et al. 2016). For this assay, the delay of 48 hr between chemical application and endpoint assessment used in the rat Cortical MEA assay minimises the chance of detecting rapid and *acute* effects of chemicals on neural network activity that would not necessarily reflect a change in neurodevelopment. Acute exposures can, however, provide information on MIEs by examining the direct effects of a chemical on cell signalling pathways that regulate neural cell electrophysiology. For example, MEA assays using mature cortical cell cultures have been used to study the effects of chemicals on action potentials following a single acute exposures (i.e., chemical addition to the cells minutes prior to recording) (Strickland et al. 2018; Valdivia et al. 2014).

27. The suitability of the chemical exposure scheme should be determined with reference to the developmental status of the test system at the beginning and end of the test for each individual test method. Information on the exposure scheme (including a graphical timeline indicating addition of chemical and medium, assessment of endpoints, etc.) in the context of the overall cell culture scheme is provided in the Section 5 of ToxTemp (Krebs et al. 2019). Further documentation of the test method should demonstrate that the exposure scheme is appropriate for examining chemical effects on a particular neurodevelopmental process based on the use of positive control compounds (Crofton et al. 2011).

3. Elements for establishing WoE

3.a. Background

28. Any WoE assessment, not only for DNT IVB data, should follow OECD GD 311 (OECD 2019) or equivalent internal guidance of regulatory authorities. WoE guidance provides a set of principles and elements that promotes a consistent and clear approach to ensure transparency and avoidance of unreasonable bias in the WoE decision processes. The overall aim of data evaluation is to determine data quality, based on relevance and reliability.

29. Relevance should be determined by an a priori problem formulation as a critical first component in the use of data from the DNT IVB in regulatory assessments. A “fit for purpose” or “context-dependent’ approach (Andersen et al. 2019; Cote et al. 2016; Griesinger et al. 2016; Judson et al. 2013) can be used to determine whether the uncertainties in the data (or dataset) are deemed acceptable for the regulatory need (see also: Parish et al., 2020; Patterson et al., 2021; van der Zalm et al., 2022; Carmichael et al., 2022) . IATAs encompass this concept (see Section 2.a and Figure 2.1) with two possible outcomes after consideration of all empirical information against the regulatory need; either the available information is adequate and supports a decision, or the uncertainty due to data gaps is too large and a decision cannot be made.

30. Reliability in a WoE will require assessments of both data for the individual assays (see Section 3.e below) as well as the entire data set from all assays (see Section 3.f below). Ideally, all assays used in the DNT IVB and the battery itself should have been validated for predicting adverse in vivo developmental neurological outcomes

according to OECD principles (see uncertainty section 4.e). However, the current lack of traditional approaches to validation should not impede the use of data from the DNT IVB in a fit-for purpose manner for some regulatory assessments (see Parish et al., 2020; Patterson et al., 2021; van der Zalm et al., 2022; Carmichael et al., 2022). The important issue is that the uncertainties in the assays and available data are transparently communicated to regulatory authorities, who will decide if such uncertainty is acceptable given their decision needs (Weinberg 1972). To date, data from the DNT IVB was used in a WoE-based decision for some organophosphate pesticides (USEPA 2020b), in IATAs for deltamethrin and flufenacet (EFSA et al. 2021), and in an IATA for neonicotinoids that include data from the DNT IVB (https://www.eu-toxrisk.eu/page/en/case-studies.php).

3.b. Generic In Vitro WoE Issues

31. All in vitro data, not just data from the DNT IVB, must meet basic criteria for use in regulatory assessments. These criteria include adherence to guidance for reporting test systems and methods, data analyses, and transparency.

32. First among these criteria is whether the in vitro data to be used was generated using OECD Test Guidelines or national equivalents. If not, documentation of the test methods should, at a minimum, adhere to OECD GD286 (OECD 2018) and GD211 (OECD 2017b). GD286 provides guidance for both development and implementation of in vitro test methods using good scientific, technical, and quality practices. The goal of GD286 is to enable test developers to establish robust in vitro methods and those utilising the in vitro method to apply the assay following the established good practice, thereby reducing uncertainties in predictions from in vitro data and assisting in the acceptance of such data by regulatory authorities. All present and future DNT IVB assays should be compliant with this guidance.

33. The rapid development of in vitro technologies has led to the recognition that in vitro test methods without applicable Test Guidelines are still useful to regulatory authorities (EFSA, 2021). OECD GD211 (OECD 2017b) outlines the type of information that should be provided for each assay, by the assay developers. The test system and methods should be described in enough detail to allow assessment of data quality and its potential use in regulatory applications. It is important to note that GD211 acknowledges that due to rapid technological developments this guidance "may thus need to be revised in the near future". Indeed, a revision has already been developed (i.e., ToxTemp) that has improved on details that should be provided for NAMs (Krebs et al. 2019).

34. The assays included in the DNT IVB proposed in this guidance were developed consistent with the principles of OECD GD 286, and all assays have descriptions that include the information required in either the original GD211, or the newly expanded ToxTemp format (Krebs et al al. 2019).

3.c. WoE Issues Specific to DNT Battery

35. There are limitations of the DNT IVB that must be considered in any WoE assessment. Some are similar to limitations of other in vitro test methods. These include lack of, or unknown metabolic competence of the assays, limitations of testing methods for volatiles or DMSO insoluble chemicals, limited exposure durations, and potency estimates based on nominal media concentrations (Thomas et al. 2018; Tice et al. 2013). Examples of limitations specific to the DNT IVB are listed in Table 3.1 (see also section 4.e).

Table 3.1. Examples of WoE limitations in evaluation of DNT IVB

- The lack of assays for several cellular processes and systemic processes known to be critical for normal neurological development (see Sections 2.b and 3.f).
- Need for development of additional AOPs to increase mapping of the KEs covered in the DNT IVB.
- A relatively limited number of tested chemicals as compared to current accepted batteries (e.g., ER activation).
- Uncertainty in the overall specificity and sensitivity of the DNT IVB due to limited testing of DNT reference chemicals and comparison of results to a curated in vivo developmental neurotoxicity study database.
- A need for a consensus-based and regulatory driven tiered testing strategy to be used in IATAs.

3. d. Data Availability

36. A key component of any WoE is transparency in all decisions. For high-throughput screening (HTS) data an especially critical component is the availability of all data on the activity or inactivity of a compound in an assay. Examples of good practice in data availability and analyses include the Tox21 and ToxCast datasets available from NCATS (https://tripod.nih.gov/tox21/assays/), NTP (CEBS database), and EPA ToxCast programs (https://www.epa.gov/chemical-research/exploring-toxcast-data-downloadable-data), in which the publicly available data is adequate to foster transparency and independent replication of analyses.

37. To aid in transparency and increasing confidence in data usage, it is critical that there is public access to both all the underlying data as well as the analysis algorithm code. Results from some previous analyses are available as processed data (Masjosthusmann et al. 2020; Frank et al. 2017; Shafer et al. 2019; Harrill et al. 2018). Processed data has been through normalisation, curve fitting and hit calls, unprocessed data has not been through these processes. Some additional assays from the NTP project have hit calls, but only summary data are available (https://doi.org/10.22427/NTP-DATA-002-00062-0001-0000-1). The recent release of ToxCastDB v3.5 provides public access to both unprocessed and processed data, the underlying analysis code, and results for all 17 assays analysed using the ToxCast data analysis pipeline (ToxCast Website). This includes hit information, curve fit comparisons, and a number of "flags" or warnings that identify by the processing pipeline have identified possible false positive or false negative findings.

3. e. Criteria for Individual Assay Evaluation

38. Evaluation and appropriate use of the data derived from NAMs such as the individual in vitro assays that comprise the DNT IVB should be based on demonstrating that they have biological and toxicological relevance (e.g., scientific rationale for use of test system, linking endpoints measured to an adverse outcome) and acceptable technical qualification (e.g., repeatability, reliability, predictive capacity) (Judson et al. 2013). As described in Section 2, the battery of assays in the DNT IVB are based on an understanding of the underlying neurobiological processes required for normal brain development, with each assay modelling a distinct KE at the cellular level in a AOP for DNT (Figure 2.2). Evaluation of the utility and performance of any specific assay will therefore be influenced by the documentation of how well the in vitro test system can recapitulate a particular KE. The assays in the DNT IVB were developed in different laboratories using multiple technologies. Thus, while the criteria used for evaluating the state of "readiness" (i.e., fit-for-purpose use) are applicable across all assays in the DNT IVB, the type and amount of information describing individual assays, including the availability, and testing of reference chemicals, will vary. Important supporting information for the evaluation and use of data from individual assays is listed in Table 3.2 (see Krebs et al, 2019 for details).

39. It is important to note that this guidance document does not provide extensive details and analyses of the assays in the DNT IVB. Instead, it highlights those factors deemed most important for assessing how and when the data may be used in a fit-for-purpose regulatory context, including screening and prioritisation of chemicals or targeted testing for single chemical hazard assessment (see also Section 1.c). A detailed description of the information and data required for evaluation of the individual DNT IVB assays is documented in Appendix B. These descriptions may use different formats including the ToxTemp format (Krebs et al. 2019) which explicitly provides the acceptance criteria for test elements needed for evaluation of assay performance. A summary of the information described in the ToxTemp is shown in Table 3.3. Other formats used (e.g., OECD 211) include similar information. Further sources of data documenting assay relevance and performance can be found in publications from the assay developers as documented in assay descriptions in Appendix B, and several reviews evaluating the readiness of DNT in vitro assays for regulatory use (Bal-Price et al. 2018; Fritsche et al. 2017;

Table 3.2 Evaluation criteria for critical elements that support use of data from individual assays in the DNT IVB.

Critical Element	Supporting Information
The test system assesses change in one or more aspects of neurodevelopmental processes	Characterisation of neurodevelopmental process in vitro including demonstration that endpoints accurately quantify developmental change in key event over time. Evaluation of the ability of measured endpoints to quantify chemical-induced changes in the key event, as well as concurrent measurement of viability/cytotoxicity.
Assay quality and repeatability .	Estimates of criteria such as signal to noise ratio, linear range, and baseline variation of the measured endpoint(s). Within laboratory repeatability of assay is demonstrated in medium- to high-throughput testing format using multi-well plates.
Chemical screening with training set	Use of training set (endpoint-specific positive controls and negative controls) to demonstrate ability of assay to test multiple chemicals in concentration-response mode and provide initial indication of predictive ability.
Data analysis and identification of a reference point	Methods for data normalisation, curve fitting for concentration-response analysis documented. Derivation and justification of a reference point identifying critical effect level based on concentration-response data.
Descriptions of chemical activity	Description of procedure used to classify chemicals as active or inactive in the assay. Determination of selective effect on neurodevelopmental process versus general effect on cytotoxicity. Supported by correct identification of positive and negative controls.

Table 3.3. Summary of information in the ToxTemp documentation format with reference to assays in the DNT IVB (Krebs et al. 2019)	
Section	**Information**
1. Overview	Descriptive title and brief abstract describing test method and endpoints used to assess key neurodevelopmental process(es) related to DNT.
2. General information	Test method name, version, related databases, depositor, and contact person.
3. Description of general features of the test system source	Supplier and source of cells. Definition and characterisation of cells, including description of type of neural cell culture (e.g., Neural Progenitor Cells, Primary Cells), species, and format (e.g., 2-D, neurosphere). Procedures for characterisation and maintenance of cells (including acceptance criteria), and differentiation of cells towards the final test system.
4. Definition of test system as used in the test method	Description of the cell culture protocol in terms of the neurodevelopmental process tested. What subpopulation of neural cells are present during the assay (e.g., subpopulations of neurons, glia)? How is the cell culture manipulated (addition of mitotic inhibitors, growth factors, etc.) during the assay? Is there endogenous metabolic capacity in the test system (e.g., CYPs)? Is there any transporter activity? How closely does the test system reflect the *in vivo* neural tissue being modelled?
5. Test method exposure scheme and endpoints	Description of the exposure scheme: how and when are cells exposed to the test compound (including a graphical timeline) in terms of the overall culture protocol (when cells are plated, medium changes, endpoint measurement)? Definition of specific endpoints measured in relationship to the neurodevelopmental process tested. Is cytotoxicity assessed concurrently? Are any reference endpoints included (e.g., cell number, protein content) to be used for normalisation? What endpoint-specific controls (positive, negative, unspecific) are used? What rules are applied to determine that the assay is performing as expected and that the results are acceptable?
6. Handling details of the test method	Description and documentation of the execution of the assay, with reference to a Standard Operating Procedure (SOP). Provides details for preparation and addition of test compounds, including definition of the concentration range. Precautions and uncertainties that can affect with the assay should be noted (e.g., compounds that are volatile or auto fluorescent, known sources of variation, need for special supplies or equipment).
7. Data management	Description of raw data format, definition and handling of outliers, and processing of raw data to summary data.
8. Prediction model and toxicological application	Which neurodevelopmental processes are modelled in the test method? What toxicological endpoints (e.g., neurite length, cell number, etc.) are used in the prediction model? How does the model classify the results in terms of toxicity (e.g., hit, no hit, borderline) and selectivity (e.g., neurotoxicity vs cytotoxicity)? What is the performance of the test method (e.g., sensitivity and specificity, z-score, etc.) and what compounds were used to make that determination? What is the application domain of the test method and how does it fit into the DNT IVB?
9. Publication/validation status	Provide a list of key publications describing the development and use of the test method. Is the test method linked to an AOP or is there information on mechanistic validation? Is it part of a formal pre-validation or validation study?
10. Test method transferability	What experience/training is required? Has the test method been transferred to another laboratory?
11. Safety, ethics, and specific requirements	Are there any specific hazards or safety requirements? Is any aspect of the test method licensed or protected by intellectual property rights?

Masjosthusmann et al. 2020). The following text provides examples and recommendations for evaluating test elements that are critical for determining assay performance and informing data interpretation.

40. *Application of the test system to assess the key neurodevelopmental processes*. Because the assays are designed to recapitulate basic processes of neurodevelopment, the ability of the test system to reproducibly undergo quantifiable developmental changes over time using a multi-well plate format is a prerequisite for assay performance. ***All assays in the DNT IVB are known to use test systems that assess neurodevelopmental processes (Fritsche et al. 2017).*** As described in Crofton et al. (Crofton et al. 2011), cell culture test systems should be characterised and their ability to accurately quantify developmental changes in an endpoint related to the KE demonstrated. Additional endpoints to concurrently assess aspects of cell health and viability should be included. This information is typically provided in peer-reviewed publications by the assay developers and can include data showing control of the culture conditions that influence the dynamic range of the endpoint. For example, published data that evaluate the KE of neurite outgrowth are available for LUHMES cells (UKN4 assay (Stiegler et al. 2011), and human neuronal cells (hN initiation assay (Harrill et al. 2011a). These publications characterised cell growth in 96-well plate format, quantified neurite outgrowth over time using automated measurement of neurite number, length, and/or area as endpoints, and simultaneously assessed cell viability. Both studies used endpoint-specific controls (see Textbox 2) to demonstrate the dynamic range of the test system and characterise concentration-related changes in neurite outgrowth.

Textbox 2.

Endpoint-Specific Controls are chemicals that are known to alter the endpoint of concern in a particular test system, also termed "endpoint-selective controls" or "mechanistic tool compounds". Within a known concentration range, they selectively alter the key neurodevelopmental event (e.g., cell proliferation, neurite outgrowth) without affecting general test system characteristics including cell viability.

41. *Assay quality and repeatability in medium- to high-throughput testing mode*. Characterisation of the utility of in vitro assays for chemical testing should include estimates of the signal (S, endpoint measurement under maximal conditions), the background (B, endpoint at baseline or solvent control conditions) and the variation between wells in a multi-well plate format. The signal-to-background ratio (S/B = mean signal / mean background) and the signal-to-noise ratio (S/N = mean signal – mean background / standard deviation of background) can be determined to provide an indication of the dynamic range of the assay. While there is no consensus on an acceptable value, larger numbers generally indicate better dynamic range. For example, in the UKN4 assay, 30,000 cells/well and a time period of 2 h and 24 h after replating were used to evaluate neurite outgrowth, which resulted in a S/N ratio greater than 50 (Stiegler et al. 2011). However, S/B and S/N ratios do not adequately take into consideration assay variability between wells. Another measure, the z' factor, has been developed that provides a measure of assay quality that takes account of both the signal window and assay variability (Zhang et al. 1999). The higher the z' value the more discriminating the assay, with a z' value of > 0.5 generally considered acceptable for complex cell-based assays like those in the DNT IVB. In the UKN4 neurite outgrowth assay example, the z' score was 0.6 (Stiegler et al. 2011). Further discussion of these criteria can be found in OECD 286 Guidance Document on Good In Vitro Method Practices (OECD 2018).

42. Variation between plates and over time (i.e., between experimental runs) is also used to evaluate repeatability. This information can be obtained by assessing endpoint response under baseline (solvent control) and treated (endpoint-specific positive control) conditions in multiple plates across independent cultures. For example, repeatability of neurite outgrowth was compared in the hN assay and the rat primary cortical initiation assay across multiple plates and cultures (Harrill et al. 2011b). The results showed that an endpoint-specific positive control (lithium chloride) consistently decreased neurite outgrowth in both assays, with a coefficient of variation (CV) of 27% in human neurons and 6% in rat primary cortical neurons. Similar data for other assays in

592 the DNT IVB can be found in the EFSA DNT publication (Masjosthusmann et al. 2020) and Appendix B. This
593 information is important for understanding the difference in sensitivity between assays***. These results illustrate***
594 ***that the assays in the DNT IVB differ in their variability, and thus will not be able to detect the same effect size***.
595 This will be an important consideration in the calculation of a reference point and formulation of the data analysis
596 algorithm (see below), especially when comparing data across the battery of assays.

597 43. *Use of a training set*. To further demonstrate the biological relevance and utility of an assay, as well as its
598 practical ability to rapidly and efficiently screen moderate numbers of chemicals, a larger set of chemicals, called
599 a *training set* (Crofton et al. 2011) should be used (see Textbox3).

Textbox 3.

A **Training Set** is a small group of chemicals including those that are known to reliably elicit a response to the specific neurodevelopmental endpoint of interest (positives), and those that do not elicit a response (negatives). Evidence for an effect, or lack thereof, should come from in vitro data using multiple test systems.

600 44. It is important to note that for the assays in the DNT IVB, there is not a standard training set for each of
601 the neurodevelopmental processes. Rather, training sets have been selected and annotated by the individual
602 assay developers, and often include several endpoint-specific controls. Examples of training sets for in vitro assays
603 of the neurodevelopmental process of neurite outgrowth are described for the hN, UKN4 and hNPC assays (Harrill
604 et al. 2011b; Krug et al. 2013; Masjosthusmann et al. 2020). It is important to note that training set chemicals are
605 specific to the neurodevelopmental endpoint being measured, and different training sets will be necessary for the
606 different assays in the DNT IVB. In the above examples, the training sets for assays of neurite outgrowth included
607 chemicals that selectively affected signalling pathways involved in neurite extension, as well as environmental
608 toxicants that have multiple or unknown modes-of-action that include known effects on neurite outgrowth (Harrill
609 et al. 2011b; Krug et al. 2013). Similarly, chemicals affecting signalling underlying hNPC migration, neuronal or
610 oligodendrocyte differentiation were used as a training set for those endpoints (Masjosthusmann et al. 2020). In
611 comparison, a training set for an assay for proliferation included chemicals that selectively altered DNA replication
612 and cell division, as well as some of the same non-selective environmental toxicants that affect neurite outgrowth
613 that also affect proliferation (Mundy et al. 2010). The use of a training set which includes chemicals that have a
614 positive effect, or no effect informs the demonstration of the ability of an assay to test multiple chemicals at the
615 same time in concentration-response mode (i.e., throughput), and can provide an initial indication of its predictive
616 ability. ***All assays in the DNT IVB were developed using appropriate training sets.***

617 45. *Data analysis and identification of a reference point*. The methods used to analyse data from the DNT IVB
618 assays require two major steps. The first is a statistical data analysis algorithm, and the second is the choice of a
619 reference point (see Textbox 4).

Textbox 4.

The **Reference Point** is a point on the concentration-response curve corresponding to an estimate of potency (e.g., EC_{50}) or a threshold defining a critical level of response (e.g., 20% change from control).

620 46. Data analyses typically start with plate corrections. Assays in the DNT IVB normally include solvent
621 controls and a wide concentration range for a chemical within a multi-well plate. This "plate map" is then repeated
622 across multiple plates, and for large numbers of chemicals testing can occur over multiple cultures and
623 experimental runs. The initial step may include "pre-processing" of raw plate data, such as subtracting out plate
624 blanks (Nimtz et al. 2019). This step may vary depending upon the technology used to measure the endpoint (e.g.,
625 fluorescence plate reader, automated imaging, etc.). The next step is normalisation of data to control wells within
626 a plate to correct for plate-to-plate variability. In high-throughput screening assays there are several methods for

normalisation based on calculating percent of control. These include dividing raw values by the average of the within plate solvent control, or, under the assumption that most compounds are inactive, including wells containing low chemical concentrations in the average of the solvent controls (Malo et al. 2006). Normalised concentration-response data are then pooled across all experiments and subjected to curve fitting using multiple models (including linear and non-linear models), and the best fit curve selected and used for determination of a reference point. ***The normalisation and curve fitting methods have been clearly stated for each assay in the DNT IVB and are found in the assay description in Appendix B.*** Automated data analysis pipelines have been developed for high-throughput screening assays (e.g., Filer et al. 2017; Hsieh et al. 2015), and models are also available to estimate the uncertainties in estimated potency and efficacy endpoints (Watt and Judson 2018). Moving forward, data analysis should be performed in a consistent and transparent manner for all assays in the DNT IVB. Confidence in any results is vastly improved when all aspects of the statistical analyses are publicly available and allow replication of findings.

47. A consensus choice of reference points (effect concentrations) for chemical effects observed using in vitro assays in the DNT IVB has not yet been formalised (Hardy et al., 2017). Based on standard quantitative analysis of chemical-receptor interactions, many in vitro pharmacologic and toxicologic studies focus on determining the concentration giving rise to a 50% response (i.e., EC_{50}) (Goodman 1996). While using this approach can be valuable in ranking the potency of chemicals and prioritising them for further testing (e.g., Paul Friedman et al. 2016), it does not necessarily consider biological variation or relevance of the degree of in vitro change to possible in vivo adverse outcomes. For this purpose, the Benchmark Concentration (BMC) approach has been recommended by the EFSA Scientific Committee (Hardy et al. 2017). This approach fits a concentration-response curve to all data points and generates the BMC as the concentration that is associated with a predefined level of response, that is, the Benchmark Response (BMR). Ideally the in vitro BMR would be chosen based on quantitative understanding of the relationship between the endpoint measured (neurodevelopmental process) and the adverse outcome (DNT), as demonstrated by a quantitative AOP. There are currently no documented quantitative relationships between DNT IVB endpoints and in vivo neurodevelopmental outcomes. Thus, without a data-derived quantitative relationship, expert judgement and scientific consensus of stakeholders can supply a suggested in vitro BMR for a particular endpoint. ***For individual assays in the DNT IVB this choice has, to date, been made by the assay developer. Based on variation in the particular assay, the current parameters used are considered adequate to limit the number of false positives.*** Ongoing work is aimed at development of a consolidated data analysis pipeline that uses the same process for choices of in vitro BMRs across assays, to begin to address this issue (Paul Friedman, personal communication). This effort will allow estimation of the false positive and false negative rates of the entire battery. Preventing false negatives in DNT IVB screening efforts is critical and is needed to decrease the uncertainty in hazard assessments.

48. In the absence of a biological basis for selecting the in vitro BMR, its value can be defined statistically as an effect size that is higher than the biological (control) variability of the measured endpoint (Davis et al. 2011). It is important to note that this statistical approach can result in different in vitro BMRs for each endpoint, since the variability will change with the different test systems and methods used in each assay. On a practical basis the in vitro BMR is determined by quantifying the variability across all control wells in all plates used within a data set or experiment. For screening studies, this can include hundreds of plates. Again, there is not a formalised approach for either calculating the assay variability or setting the in vitro BMR using this calculated value, and the decision is left to the assay operator. One method used in evaluating variability of high-throughput screening data in the USEPA ToxCast program is determining the baseline median absolute deviation (BMAD), a robust statistic of control variability that is resilient to outliers (Leys et al. 2013). This approach yielded coefficients of variation in control values ranging from 2-22% across some in vitro DNT assays (including the endpoint of cytotoxicity) developed at the USEPA (Harrill et al. 2018). Another method normalises the lowest compound concentration (assuming the lowest concentration has no affect) to the solvent controls within a plate and calculates the standard deviation (SD) between the means of the lowest concentration over all plates in an experiment. This

resulted in deviations ranging from 1.5-27% across the assays developed at the IUF - Leibniz Research Institute for Environmental Medicine (Masjosthusmann et al. 2020). Once the variation has been determined, a modifying factor is used to account for the distribution of values in the control well population. Thus, the in vitro BMR may be set for example, at a level above 2xSD (accounting for 95% of the variation in control wells) or 3xSD (accounting for 99% of the control wells).

49. Currently, in vitro BMRs have been set for some assays based on both expert judgement and statistical considerations as determined by the assay developer. For example, for cell migration (UKN2 assay), in vitro BMRs of 25% for migration and 10% for cytotoxicity were used based in part on expert judgement of in vitro biological significance (i.e., what is a meaningful extent of reduction), in part on statistical considerations (all positive controls showed a reduction in migration of > 25%), and in part on graphical comparisons of effective concentrations for positive control chemicals and "unspecific" controls (chemicals that show general cytotoxicity in cell cultures) (Nyffeler et al. 2017b; Zimmer et al. 2012).

50. *Data Analysis and Statistical Model*. Data analysis for the in vitro DNT assay results requires a statistical model which is an algorithm or set of rules for determining whether a chemical has altered the key neurodevelopmental process assessed in the assay. These decisions should include: 1) an evaluation as to whether a change in an endpoint has reached or surpassed the level set as the in vitro BMR; and 2) a determination as to whether the effect is *selective* for the neurodevelopmental process examined or is a result of a general effect on cell health and viability (i.e., distinguish "DNT-specific hits" from "nonspecific hits"). The former is based on the statistical analyses of the data to determine whether a threshold for a hit has been met or exceeded. To determine selectivity, the chemical concentration affecting the neurodevelopmental endpoint (e.g., migration distance, neurite length, etc.) is compared to a concurrent measurement of the concentration affecting cell health (e.g., a measure of cytotoxicity or viability).

51. The statistical algorithm would then categorise a chemical as either *inactive*, *active and selective*, or *active but non-selective (Textbox 5)*. In some cases, a fourth category defined as *borderline* has been used (Delp et al. 2018; Masjosthusmann et al. 2020). Borderline classifications may be defined by an overlap of the confidence intervals for the in vitro BMC of the neurodevelopmental process and in vitro BMC for cytotoxicity/viability (Masjosthusmann et al. 2020). Examples for chemical concentration-response curves illustrating each category have been reported (Delp et al. 2018; Harrill et al. 2018) and are illustrated in Figure 22 from Masjosthusmann et al. (Masjosthusmann et al. 2020).

Textbox 5. Potential Chemical Categories Based on Assay Results:

Inactive – An in vitro BMR not reached for either the neurodevelopmental process or cytotoxicity/viability

Active and selective – An in vitro BMC for the neurodevelopmental process separated from the BMC for cytotoxicity/viability

Active but non-selective – An in vitro BMC for the neurodevelopmental endpoint is not separated from the BMC for cytotoxicity/viability

Borderline – chemicals for which the separation between the in vitro BMC for the neurodevelopmental endpoint and the in vitro BMC for cytotoxicity/viability is not clear

52. From the description provided above it is apparent that the decision as to whether a chemical has a selective effect on neurodevelopment (i.e., is a DNT-specific hit) will depend upon the choice of the level of response set for the in vitro BMRs for the neurodevelopmental endpoint and a non-specific endpoint such as cytotoxicity/viability, and the procedure used to define the degree of separation required between the resulting in vitro BMCs. To date, the procedure for determining the degree of separation is not standardised but has been set by the assay developer. For the DNT IVB assays, a common approach has been used whereby testing of each assay includes both *endpoint-specific controls* (positive control chemicals that alter the endpoint of interest) and *unspecific controls* (chemicals that show general cytotoxicity in cell cultures). The degree of separation of the in

711 vitro BMCs for the neurodevelopmental endpoint and cytotoxicity is compared for each group of controls, with
712 the expectation is that it will be larger for the endpoint-specific controls and smaller for the unspecific controls.
713 Another approach is to use cytotoxicity data from multiple assays to determine the cytotoxicity "burst" (c.f.,
714 Judson et al., 2016; Escher et al., 2020). All these approaches assume that the presence of cytotoxicity is not just
715 a confound, but instead a cause of any changes in the assay endpoint. The degree of separation between the
716 concentrations that change the DNT assay and the cytotoxicity assay will impact the rate of false positive and false
717 negatives, and the criteria can also be adjusted to balance or limit false negatives, should this be a priority for
718 regulatory use.

719 53. Statistical determination of the degree of separation most often examines the ratio of the in vitro BMC
720 for cytotoxicity to the in vitro BMC for the neurodevelopmental endpoint for each unspecific control chemical
721 (e.g., EC_{50} cytotoxicity/EC_{50} neurodevelopmental endpoint). The ratios are averaged and the variation (SD or 95%
722 confidence interval) calculated. The degree of separation is then chosen as a value that accounts for the upper
723 bound of the variation of the ratios. For example, in the rat cortical MEA Neural Network Formation assay, the
724 mean of these ratios was 1.9 (i.e., network formation endpoints were on average affected by unspecific controls
725 at slightly lower concentrations than cytotoxicity). The upper 95% confidence interval of the ratio was 2.8, so a
726 ratio of 3 was chosen as a threshold for chemicals to be considered as having selective effects (Frank et al. 2017;
727 Shafer et al. 2019). Similar calculations were done for neurite outgrowth in the UKN4 assay, resulting in an average
728 ratio of 1.4 with a SD of 0.8 (Krug et al. 2013). In this case, variation was accounted for using 3xSD, and a ratio of
729 4 was used as the threshold (see Krug et al. 2013; Delp et al. 2018). To verify the appropriateness of the threshold,
730 ratios from the endpoint-specific controls and unspecific controls should be examined to see if they are classified
731 correctly as selective or nonselective, respectively.

732 54. A second approach is to set individual fixed in vitro BMR levels for the neurodevelopmental and
733 cytotoxicity/viability endpoints which consider differences in the underlying biological process and/or the baseline
734 variability. Different in vitro BMR levels can be compared, and the most appropriate chosen based on a ratio that
735 correctly classifies the endpoint-specific and unspecific control chemicals. This approach has been used for the
736 UKN2 cell migration assay (Nyffeler et al. 2017b). An in vitro BMR of 10% for cytotoxicity/viability was used based
737 on biological plausibility that changes in viability below this value are not meaningful. In contrast, an in vitro BMR
738 of 25% for migration was used based on experimental findings that unspecific control chemicals may cause up to
739 a 25% change of migration, but not beyond. Using the EC_{10} cytotoxicity/EC_{25} migration ratio, all unspecific control
740 chemicals had a ratio ≤ 1.11, whereas endpoint-specific control chemicals reached a ratio of > 1.3. Thus, the EC_{10}
741 cytotoxicity/EC_{25} migration ratio of 1.3 was used as the threshold to classify selective chemicals (Nyffeler et al.
742 2017a).

743 55. As described above, the outcome of the data analyses that interpret chemical data from individual assays
744 in the DNT IVB is dependent upon variables that are inherent to the test method (complexity of the test system,
745 measurement technology used, biological variability, etc.), but also upon decisions by the assay
746 developer/operator. These decisions include expert judgement on the level of change in an endpoint considered
747 as biologically relevant in vitro (i.e., relative to assay positive controls in the training set), what constitutes a "DNT-
748 specific" effect, and the statistical evaluation of variability. Currently, most statistical algorithm set parameters
749 that are relatively conservative and maximise both sensitivity (the correct prediction of a positive compound) and
750 specificity (the correct prediction of a negative compound) (e.g., Filer et al. 2017; Hsieh et al. 2015). ***These***
751 ***"default" parameters used by assay developers for assays in the DNT IVB are considered adequate, and thus***
752 ***limit the number of false positives in the individual assays.*** An important advantage of public access to the data
753 and analysis algorithms is that it allows reanalyses that can incorporate adjustment of these parameters to focus
754 on balancing or limiting false negatives, should this be a priority for regulatory use. However, screening of
755 chemicals for potential developmental neurotoxicity involves multiple (unknown) targets and complex biological
756 processes. In this case, the analysis parameters can be altered to be fit-for-purpose. When screening for

prioritisation the specificity could be altered to allow for a higher number of false positives, so that a smaller number of potential neurotoxicants will be missed. For example, in the UKN2 cell migration assay the in vitro BMR level for the migration endpoint could be reduced from 25% to 20% (leading to more chemicals classified as active) and the selectivity threshold could be reduced from 1.3 to 1.2 (leading to more chemicals classified as selective). This approach may be desirable when testing large numbers of chemicals for which there is little or no toxicology data.

3f. Evaluation of the DNT IVB for chemical testing

56. *Predictive performance.* To date there has been limited use of the current DNT IVB in chemical testing. This is mainly the result of research to date that has focused on assay development and refinement efforts. In addition, chemical testing will be facilitated by the assembly of a set of annotated *DNT reference chemicals (see Textbox 6)*.

57. The importance of a set of reference chemicals cannot be understated because it allows estimations of performance, in terms of sensitivity and specificity, of both individual assays and the assay battery. Currently there is no consensus set of reference chemicals for use in development and validation of DNT in vitro assays. Ideally, 100 or more structurally diverse chemicals should be identified. Testing this set of DNT reference chemicals across all the assays would allow for the evaluation of the predictive performance of the entire battery. Several groups have evaluated the available evidence for in vivo developmental neurotoxicants to be included as reference

Textbox 6.
DNT Reference Chemicals are chemicals with evidence for in vivo developmental neurotoxicity identified in studies of humans or animals (positives), as well as chemicals with evidence that they do not result in developmental neurotoxicity in vivo (negatives).

chemicals. Short lists of so called “gold standard” chemicals that are generally acknowledged to be developmental neurotoxicants in humans have been proposed (Grandjean and Landrigan 2006; 2014; Rees et al. 1990). Mundy et al. (Mundy et al. 2015) evaluated approximately 500 peer-reviewed publications and EPA DNT guideline studies and found evidence for in vivo developmental neurotoxicity in mammals and/or humans for approximately 100 chemicals. Importantly, this list included only chemicals where empirical evidence of DNT was reported from two or more different laboratories. Similarly, a workshop consensus (Aschner et al. 2017) identified approximately 30 chemicals as being in vivo developmental neurotoxicants. Individual laboratories employing assays in the DNT IVB have, to date, used lists of proposed positive compounds derived from the above publications that include empirical findings of in vivo DNT. Only recently has there been a review publication for negative reference compounds, and it lists only eight chemicals that have empirical findings of no DNT in vivo (Martin et al., 2021). Thus, in the work examining DNT IVB assays, these same laboratories used more extensive lists of “proposed” negative reference chemicals, chosen primarily based on the authors’ expert opinions and not necessarily on empirical negative in vivo data (see in publications listed below for details). Performance estimates of the DNT IVB assays, indicated by the ability to correctly detect a positive (sensitivity), and reject a negative (specificity) abased on positive and negative reference chemicals chosen by the authors are summarised below. A summary of all tested reference chemicals is in Appendix A.

58. A set of 75 proposed DNT reference chemicals was assessed in the cortical MEA network formation assay (see Figure 2.3) developed at the USEPA (Frank et al. 2017; Shafer et al. 2019). The chemical set included 61 proposed positives and 14 proposed negatives. Based on a selective hit (i.e., a hit that is DNT endpoint-specific as compared to cytotoxic) in at least one of the seventeen network parameters assessed, the sensitivity was 61% and the specificity was 86%. If the data is considered in terms of any active hit (selective or not selective) in at least one network parameter, the sensitivity was 75% and the specificity was 86%.

59. A set of chemicals consisting of 53 proposed positives and 14 proposed negatives were tested in a suite of DNT IVB assays (see EPA assays in Figure 2.3) including hNP1 proliferation, hNP1 apoptosis, hN neurite initiation, cortical neurite initiation, and cortical maturation and synaptogenesis (Harrill et al. 2018). Based on selective effects, the combined assays had a sensitivity of 68% and specificity of 93%. When all active hits are considered (regardless of selectivity) the sensitivity was 87% and the specificity was 71%. In both cases, the combined results of all five assays had a better sensitivity and specificity than any individual assay (Harrill et al. 2018).

60. Output from the European Food Safety Authority (EFSA) and Danish EPA projects involved testing at the University of Konstanz (UKON) and the IUF in Düsseldorf using ten DNT IVB assays (see Fig 2.3) This work tested a library of about 100 compounds, assembled from multiple sources, including 29 proposed DNT positives and 17 proposed DNT negatives (for details see Masjosthusmann et al. 2020). The assays tested endpoints including, proliferation, migration, differentiation, neurite outgrowth, and oligodendrocyte differentiation. For selective effects, the combined assays had a sensitivity of 83% and a specificity of 88%. Based on all active hits (regardless of selectivity) from the combined assays, the sensitivity was 83% and the specificity was 82%. Further analysis showed that performance was optimal when results from all assays were combined (Masjosthusmann et al. 2020).

61. The currently available data indicate that use of all 17 assays currently in the DNT IVB can detect a majority of chemicals considered to be in vivo developmental neurotoxicants, with a relatively small number of false positives (Masjothusmann et al., 2020; Harrill et al., 2018; Shafer et al., 2019). The data support the previous international consensus that the battery could be used in the current form for chemical screening (Fritsche et al. 2017; OECD 2017a). As discussed below, a more complete understanding of the predictive performance of the DNT IVB will be possible when a standardised DNT reference chemical set is tested using the entire suite of assays (Table 2.3). There is currently an ongoing effort to re-analyse all data from both EPA and EFSA projects using a common analysis pipeline, an effort likely to be completed in 2022 (Paul Friedman, personal communication). ***Such analyses will be critical to estimate performance of the entire DNT IVB, and importantly, allow interlaboratory comparisons for assays that measure similar processes.***

62. ***Based on selective hits, the predictive ability across all assays in the DNT IVB exhibit a sensitivity range from 61 to 83%***. This is comparable to many other in vitro screening assays. For context, performance estimates of the DNT IVB, can be compared to in vitro assays used for chemical screening for other outcomes such as carcinogenicity, hepatotoxicity, and endocrine disruption. The ability of a battery of commonly used in vitro genotoxicity tests (Ames assay, mouse lymphoma assay, and the in vitro micronucleus or chromosomal aberrations assay) was evaluated for its ability to predict rodent carcinogens and non-carcinogens, based on a database of over 700 reference chemicals (Kirkland et al. 2005). Using an optimal combination of three assays the sensitivity was 93%, but the specificity was only 29%. Vorrink et al. (Vorrink et al. 2018) evaluated the use of human hepatic spheroid cultures to predict hepatotoxicity of 123 reference compounds with positive or negative clinical evidence for drug-induced liver injury. The results showed a sensitivity of 69% and a specificity of 100%. Judson et al. (2015) used a battery of assays to screen for oestrogen receptor active chemicals with a set of 18 in vitro assays covering multiple KEs. Computational model classifications were derived for 1812 substances, and model performance was compared to data for specific reference chemicals. Depending on the combination of tests used, the prediction model achieved a sensitivity of 97% and specificity of 89%, as compared to guideline in vitro tests and the uterotrophic assay.

63. Two factors contribute to the uncertainty in assessing performance of the battery. The first is the lack of a comprehensive analysis to assess performance of the entire data set. The second is the extent of coverage in the DNT IVB for all critical neurodevelopmental processes (Figure 2.3). Several gaps in coverage of neurodevelopment processes and cell types have been acknowledged, including assays for neuroectodermal formation, astrocyte differentiation and maturation, the blood-brain and placental barriers, microglia regulation of neuronal growth and connectivity, neuronal subtype specification, and axon myelination (see Section 2). It is

expected that addition of assays that account for these aspects of neurodevelopmental processes will increase the predictive ability of the battery.

64. Several factors, also common to most in vitro assays, contribute to the uncertainty in predicting in vivo developmental neurotoxicity. Brief descriptions of the most important of these are listed below. For a more complete review see Barbosa et al. 2015; Fritsche et al. 2015; Harry et al. 1998; Harry and Tiffany-Castiglioni 2005; Lein et al. 2005; Smirnova et al. 2014.

- Incomplete coverage of complex interactions within the brain during development. Factors that regulate normal brain development in vivo including cell-to-cell communication within and between brain regions, neurotransmitter and growth factor signalling (Cameron et al. 1998; Cowan and Petri 2018; Ojeda and Avila 2019) are not fully accounted for in the current neural cell culture models. Multicellular test systems (including neurospheres and rat primary cortical cultures) contain limited aspects of cell-to-cell interactions like auto- or paracrine signalling (Kartvelishvily et al. 2006; Masjosthusmann et al. 2018; Ogunshola et al. 2002), thus there is a need for additional assays to cover these complex interactions.
- Use of human- and animal-derived neural cell cultures. Some assays in the DNT IVB use animal derived cells rather than human cells, and the potential for species-specific effects is unknown. Also, the current DNT IVB does not fully account for sex or human genetic diversity that may influence susceptibility to chemical-induced developmental neurotoxicity (i.e., gene x environment interaction). These factors may result in lower sensitivity and specificity.
- ADME (absorption, distribution, metabolism, and excretion). Like most other in vitro assays/batteries, the ADME processes that determine chemical exposure in vivo are mostly absent in the current DNT IVB. Many in vitro test systems, including the DNT IVB, have minimal or unknown metabolic capacities compared to liver (DeGroot et al. 2018; Ferguson and Tyndale 2011; Hedlund et al. 2001). In addition, developmental changes in in situ metabolism including the influence of glia are not completely understood. Thus, the DNT IVB may not accurately predict the potential DNT of chemicals that are activated or detoxified. Importantly, there are issues specific to nervous system exposure during development that are not currently included in the DNT IVB, e.g., test methods and kinetics models for chemical transport across the placental (Wong et al., 2020; Gingrich et al., 2021) and blood brain barriers (Ball et al. 2013; Barbosa et al. 2015; Delsing et al. 2020).
- Systemic impacts of hormones and immune signalling on brain development. Circulating steroid and thyroid hormones are known to impact a wide variety of neurodevelopmental processes including sexual dimorphism (Adhya et al. 2018; Bernal 2022; Arambula et al., 2020). Chemicals that disrupt maternal thyroid hormone levels in vivo can result in developmental neurotoxicity (Miller et al. 2009;Zoeller and Rovet 2004). Several in vitro screening assays have been employed to test the ToxCast and/or Tox21 chemical libraries for a number of known MIEs linked to disruption of thyroid homeostasis and consequent downstream impacts on the developing nervous system (Hornung et al. 2018; Olker et al. 2019; Paul-Friedman et al. 2019; Paul Friedman et al. 2016; Wang et al. 2018). It should be noted that some assays in the DNT battery may be modified to study hormonal mechanisms (e.g., Klose et al., 2021)., but they do not currently meet the four criteria for inclusion in the batter In addition, immune regulators (microglia), the gut/brain axis, and the placental/brain axis also regulate brain development (Cowan and Petri 2018). While endocrine activity and immune signalling are outside the domain of the DNT IVB, incorporation of such data streams will expand coverage of MIEs and KEs for known DNT AOPs,

and thus reduce the probability of false negatives. This is an example of how results from the DNT IVB should be assessed in the context of all other available relevant data.

65. The next step to a better understanding of the predictive performance of the entire DNT IVB requires a thorough analyses of all data from all 17 assays. The recent release of ToxCastDB v.3.5 has allowed this process to begin. The combined in vitro data can then be compared to any existing in vivo findings, including compounds in a consensus DNT reference chemical list (see Section 80). At this time there are no predictive in silico computational DNT models (Crofton et al., 2022). Development of computational models that use data from multiple assays of the DNT IVB for prediction could be feasible, similar to the model describing oestrogen receptor activation by Judson et al (2015), who used data from in vitro assays covering multiple KEs of the oestrogen receptor activation pathway to construct a computational model predicting endocrine disruption. This approach requires the ability to link measured MIEs, KEs and adverse outcomes, but current AOPs describing developmental neurotoxicity are incomplete. However, information on important signalling pathways modulating neurodevelopment is available for some assays in the DNT IVB (Masjosthusmann et al. 2020; Sachana et al. 2021). These data, obtained by using selective pathway inhibitors, begin to describe the biological applicability domain of the DNT IVB, and provides the basis for further AOP development. Until these types of analyses are completed, it is recommended that all available assays in the DNT IVB should be used within a fit-for-purpose approach.

66. *Evaluation of results across the battery.* Data from chemical testing using all 17 assays in the DNT IVB summarised above (Frank et al. 2017; Harrill et al. 2018; Masjosthusmann et al. 2020; Shafer et al. 2019) show that most chemicals that are hits are active in more than one assay. This is not surprising, and likely the result of two reasons. First, the key neurodevelopmental events modelled in the in vitro battery are controlled by signalling pathways that are not specific for individual processes, but rather contribute to multiple processes (Masjosthusmann et al. 2020). Thus, chemical actions at a single target (MIE) can affect multiple downstream KEs and neurodevelopmental endpoints. Second, some known developmental neurotoxicants are pleotropic, i.e. may act at multiple MIEs (e.g., Klocke and Lein 2020; Prince et al. 2019). As a result, chemical testing has shown a spectrum of chemical effects across the in vitro battery.

67. Evaluation of the potency and selectivity of chemicals across all assays in the battery may be useful in some regulatory frameworks for chemical prioritisation and may provide information on chemical hazard, as well as in the context of cumulative risk assessment Evaluation of results across the DNT IVB will be facilitated using a common set of methods and parameters to analyse data obtained from different assays and laboratories. This would include both the concentration-response analysis and determination of a reference point and selectivity (i.e., a DNT-specific hit). Currently the parameters summarising chemical response data from the individual assays (in vitro BMRs and BMCs, selectivity, etc.) have been derived using a variety of different approaches (see Section 3.e above). It is suggested that a consensus should be sought for the definition and statistical determination of the values used as the BMR level and the degree of separation between the DNT-specific endpoint and cytotoxicity (i.e., selectivity). There is currently an ongoing effort to harmonize this process (Paul Friedman, personal communication).

68. As potency is a critical determinant of hazard, chemicals can be ranked and prioritised based on comparison of relative potencies across all assays in the battery. For example, chemicals with potent effects in multiple assays could be ranked higher compared to chemicals that are potent in just one assay, followed by chemicals with lower potencies. Further development of a consensus-based statistical algorithm is needed for ranking and prioritizing results from the DNT IVB. Regardless, several procedures and tools have been applied to the data currently available for the DNT IVB in order to provide both an overall ranking of chemicals and a graphical visualisation of chemical groups or clusters with similar properties.

- Selective versus non-selective effects. The use of selectivity to filter in vitro data will exclude chemicals that are designated as active as the result of a non-specific effect on cell health. Current DNT IVB data

928 described above suggest that accounting for cytotoxicity improves the correct rejection of chemicals
929 that are DNT negatives (i.e. specificity) but may decrease identification of DNT positives (i.e., sensitivity)
930 (e.g., Harrill et al. 2018). Some well-known DNT chemicals (e.g., methylmercury) are very potent in vitro,
931 and the degree of separation between the DNT endpoint-specific in vitro BMCs and cytotoxic BMCs can
932 be small; an appropriate cut-off for selective effects can be difficult to define. As more data is collected
933 on chemical effects, it will be possible to compare results from the battery both with and without using
934 selectivity as a filter. In addition, potency estimates for both DNT endpoint-specific effects and
935 cytotoxicity can be compared to values obtained for in vitro assays examining other (non-DNT) effects
936 (see section 3g).
937 • MSE (Most Sensitive Endpoint). As an initial approach, a chemical can be ranked by potency using the
938 lowest in vitro BMC identified after testing in the DNT IVB. The assay with the lowest in vitro BMC is
939 designated as the MSE, and in addition to use for ranking, it may be considered as a relevant endpoint
940 for further targeted testing. There may be other assays, however, with in vitro BMCs in the range of the
941 MSE that should also be considered. Ranking chemicals by the MSE is illustrated in Masjosthusmann et
942 al. (2020). A caveat on the use of the MSE is that it compares assays that use different benchmark
943 response levels (e.g., BMR of 10% for radial glia migration (NPC2a) versus an in vitro BMR of 30% for
944 oligodendrocyte differentiation (NPC5). In theory this could bias the MSE to assays with lower
945 benchmark responses.
946 • Hierarchical cluster analysis. Chemical effects based on both the number of endpoints affected and the
947 potency for each endpoint can be analysed using hierarchical clustering. The result is a heatmap that
948 provides a grouping of chemicals with similar effect profiles (e.g., Harrill et al. 2018; Masjosthusmann et
949 al. 2020). The data can be used to prioritise chemicals based on potency and number of endpoints
950 affected and facilitate identification of chemicals with a common mode-of-action.
951 • ToxPi (Toxicological Prioritization Index). ToxPi is a data modelling tool that combines multiple sources
952 of data and can provide prioritised orders of chemical toxicity based on endpoint classes, as well as a
953 visualisation of the weight of the contributing factors (Reif et al. 2010). The use of ToxPi to compare
954 chemicals in a class and rank their potential for developmental neurotoxicity is illustrated in
955 Masjosthusmann et al. (2020).
956

957 69. *Consistency of evidence within the DNT IVB.* Examination of chemical outcomes across the DNT IVB can
958 also be considered in terms of consistency between multiple assays for the same neurodevelopmental endpoint
959 (e.g., multiple assays for neurite outgrowth) or known biological relationship between assays at increasing levels
960 of complexity (e.g., the upstream KE of neurite outgrowth proceeding and necessary for the downstream KE of
961 neuronal network formation). In the first case, quantitative analysis has shown that chemical effects in related
962 assays (e.g., UKN4 and UKN5 assays for neurite outgrowth) can be highly correlated (Masjosthusmann et al. 2020).
963 For an individual chemical, consistency of an outcome between different assays for the same endpoint increases
964 the strength of evidence. In the second case, data showing that a chemical alters two KEs in a neurodevelopmental
965 pathway (e.g., neurite outgrowth in the cortical maturation assay and neuronal network formation in the cortical
966 MEA assay) increases the biological plausibility of the effect.

967
968 ***3g. Chemical potency in DNT IVB assays versus assays for other endpoints***
969

70. An important consideration in the evaluation of findings from the DNT IVB is the comparison of chemical potency in neurodevelopmental assays to potency in other in vitro toxicity assay. This allows a differentiation between neural specific effects and disruption of non-specific cellular processes such as cell stress. This can be considered in two ways.

71. The first approach is comparing the potency of a given chemical that is considered a DNT-specific hit to data from other cell types., i.e. non-neural cell types. To do this the potency of DNT-specific chemicals from the DNT IVB could be compared to their potencies calculated in the "burst" region from all other non-DNT in vitro assays(Judson et al. 2016). This burst region is the concentration range showing in vitro activity across multiple assay endpoints that is associated with non-specific cellular processes. Examples include activation of cell stress pathways, disruption of proteins or membranes, or broad low-affinity non-covalent interactions (Judson et al. 2016). This approach, while not conducted yet for DNT IVB results, can inform the specificity of the DNT effect relative to non-specific effects across a wide biological spectrum.

72. The second approach is to compare the potency of a DNT-specific hit to the potency in specific hits from other non-DNT assays (e.g., assays for hepatic nuclear receptors, hormone transactivation, proliferation, apoptosis) (Thomas et al. 2013). This has been done to a limited degree for some assays in DNT IVB (Delp et al. 2018; Masjosthusmann et al. 2020; Shafer et al. 2019; Ryan et al., 2016). These comparisons have demonstrated that for some chemicals the in vitro potency estimates for the DNT IVB assays were below those found for all Tox21 assays (e.g., Leist graph reference), demonstrating the importance of DNT in vitro assays as part of any overall assessment. Failure to include DNT in vitro assays would, for some chemicals, underestimate in vitro potency.

4. Integration of Evidence

4.a. Background

73. Use of evidence from the DNT IVB may be guided by several factors, including: 1) the consistency of in vitro data derived from complementary assays within the battery itself; 2) biological plausibility based on existing AOPs for adverse developmental neurological outcomes and any available in vivo data; 3) incorporation of available in vitro to in vivo extrapolation (IVIVE) exposure modelling that extrapolates in vitro exposures to in vivo; and 4) weighting of known uncertainties of the IVB and any existing in vivo data against the regulatory needs. These factors are all integrated with the IATA framework.

4.b. Predictive Power

74. The predictive power of the battery is currently good (See Section 3.b.); estimates of sensitivity range from 61 to 87%, and specificity from 71 to 93% for different groups of DNT assays (Harrill et al. 2018; Masjosthusmann et al. 2020; Shafer et al. 2019)." These numbers will evolve, likely improve, over time as more chemicals are tested, additional negative controls are tested, new assays are developed that fill known biological gaps in the current in vitro models of neurodevelopmental processes, and also with the development of consensus-based data pipelines and publicly available databases.

4.c. Plausibility

75. Any use of results from the DNT IVB must be integrated with all other available information to determine the potential plausibility of xenobiotics to cause alterations in the developing nervous system. Data integration should be conducted using WoE methods such as OECD GD 311 (OECD 2019) or equivalent internal guidance of regulatory authorities. This may include data from in silico or read-across predictions, data from in vitro assays or alternative species, data from in vivo animal studies, and human data from clinical or epidemiological studies. Evaluation of findings for single chemicals, or structurally similar chemicals, should be made in the context of known neurobiology pathways that underlie neurodevelopment. Note that for some chemicals or chemical classes (e.g., new or untested chemicals) data on DNT may be non-existent or extremely limited. Below are several examples of the integration of DNT IVB findings with other information that support plausibility.

76. The scientific basis of the DNT IVB is the use of assays for KEs at the cellular level that are plausibly related to the mode of action of developmental neurotoxicants in vivo. At a basic level, plausibility is supported by data demonstrating the correlation of effects of chemicals on neurodevelopmental processes in vitro and in vivo. A number of known developmental neurotoxicants alter Kes in vitro and have the analogous effect after in vivo exposures, with subsequent changes in brain structure and function (Guo et al. 2013; Jones et al. 1996; Miller 1986; Tingling et al. 2013; Yang et al. 2009; Yu et al. 2010). As a detailed example, MeHg alters the rate of radial glial migration in the NPC2 assay and neural crest cell migration in the UKN2 assay (Nyffeler et al. 2017a), as well as neuronal differentiation and maturation in the NPC3, NPC4, UKN4 and UKN5 assays (Masjosthusmann et al. 2020). And it is known that in vivo developmental exposure to MeHg disturbs neurodevelopment by altering cortical migration processes and results in adverse neuronal differentiation and cortical organisation (e.g., Choi 1986; Guo et al. 2013). Consistency with other in vitro data for the same key neurodevelopmental process also adds to the plausibility. MeHg has also been shown to alter migration in other complementary in vitro assays (Kunimoto and Suzuki 1997; Sass et al. 2001), as well as trophoblast cells (Tucker and Nowak 2018). Thus, consistency of a change in a KE in the DNT IVB assay to an upstream mechanism for the same chemical increase plausibility.

77. There are many signalling pathways that are known to initiate and regulate the Kes controlling neurodevelopment, and chemical or genetic perturbation of these pathways can result in adverse outcomes in brain structure and function. Demonstration of a chemical effect on both a key neurodevelopmental process (measured in the DNT IVB) and on a signalling pathway underlying that process provides evidence for a plausible mode-of-action. For example, distinct from its role as an esterase to hydrolyse acetylcholine in the adult brain, acetylcholinesterase (AChE) acts as a morphogen during early brain development and can regulate neurite outgrowth (Bigbee et al. 1999). Cholinesterase inhibitors that bind to the morphogenic site on AChE, including chlorpyrifos and its active metabolite chlorpyrifos oxon, can inhibit neurite outgrowth in vitro (Masjosthusmann et al. 2020), in some cases at concentrations that do not inhibit the esterase activity (Das and Barone 1999; Howard et al. 2005). Additional work showed that chlorpyrifos inhibited axon growth in vivo in developing zebrafish (Yang et al. 2011). There is also limited data suggesting chlorpyrifos can alter neurite outgrowth in vivo in mammals (Qiao et al. 2003).This data supports the plausibility of DNT IVB data for inhibition of neurite outgrowth for chlorpyrifos, and although incomplete, provides a plausible AOP leading from AChE binding to altered brain growth (USEPA 2012).

78. Plausibility is also enhanced by mapping changes in the Kes measured in the DNT IVB to existing AOPs. Appendix C maps DNT IVB assay endpoints to Kes in currently available DNT AOPs. For example, neural network formation (NNF) is a KE in two proposed AOPs, and NNF is a KE in the current DNT IVB. Recently, EFSA (EFSA et al. 2021) developed an IATA case study for the integration of all available in vitro data, including data from the DNT IVB, in a developmental neurotoxicity hazard characterization for deltamethrin, a pyrethroid insecticide

(Appendix D). This case study utilised data streams that included: data from the DNT IVB, and results of a systematic review both human and animal studies in the peer reviewed literature and DNT OECD TG426 guideline study. This was followed by development of an AOP that links changes in voltage gated sodium channels, through downstream impacts on NNF, to alterations in neurodevelopment expressed as adverse impacts on neurobehaviour. A Bayesian network analysis was used to determine the probability of occurrence of downstream Kes and identified the largest uncertainty data gap as being the lack of empirical support for the biological understanding of the key event relationship (KER) between changes in neuronal network functioning and the adverse behavioural outcome (EFSA et al., 2021). This is an example of how DNT IVB data can inform a hazard characterisation for a pesticidal chemical.

4.d. Incorporation of IVIVE.

79. As with all in vitro data, the use of data from the DNT IVB in vitro assays to inform human hazard assessment requires the extrapolation of dose to identify relevant in vivo exposure levels. Thus, it is important to understand that dose, per se, in many in vitro assays is estimated as the nominal media concentration. The relationship between the media concentration and cellular or target concentration is influenced by several variables, including, but not limited to binding to media serum proteins and/or the plastics of cell plates, active vs passive uptake into cells or multi-cellular tissues, cell density, and the degree of any metabolism and parent compound stability in the culture (Paini et al. 2019). New models are available to predict partitioning to in vitro compartments, i.e. cells, serum constituents in exposure medium, microtiter plate plastic, headspace and extracellular matrices (Kramer et al. 2015; Proenca et al. 2021). Use of media concentrations as a surrogate for cellular concentrations has, especially for lipophilic compounds, been shown to underpredict cell concentrations by orders of magnitude (Croom et al. 2015; Mundy et al. 2004; Schreiber et al. 2010; Shafer and Hughes 2010). Important to note is that actual analytical measurement of media and cellular concentrations in in vitro DNT studies, including the DNT IVB, are very rarely done. However, new computational models are available for extrapolation from in vitro to in vivo, and these are applicable to DNT IVB assays. These models involve the use of empirical data for a limited number of in vitro ADME parameters (e.g., binding to serum proteins, disappearance rates of parent chemical rates in hepatic cultures) to predict a human oral equivalent dose, a dose which would result in steady-state in vivo blood concentrations equivalent to the in vitro concentration that alters a response by 50% (e.g., Wetmore et al. 2012). Data have been developed to predict oral equivalents for thousands of chemicals (e.g., Breen et al., 2021; Sipes et al., 2017; Wetmore et al., 2012). To date, this has been done for some DNT IVB assays (e.g., Algharably et al. 2021; Shafer et al. 2019). It is important to point out that some extrapolated daily human equivalent doses have extremely high uncertainty (e.g., Wetmore et al., 2012; Wambaugh et al., 2019).. Data gaps in IVIVE for developmental toxicity, including DNT, include a need to incorporate placental transfer, and in addition estimating brain concentrations and the foetal and postnatal development of the blood brain barrier. Oral equivalents calculated from in vitro assays can be combined with human exposure estimates for large numbers of chemicals to develop risk-based prioritisations (Wambaugh et al. 2019; Beal et al., 2022). IVIVE can also be used for single chemical assessments. This was recently done to determine the 'dose relevance' of data from in vitro DNT IVB assays to in vivo rodent developmental neurotoxicity data and relevant in vivo kinetics models for development (EFSA et al. 2021). While it is acknowledged that progress has been made on development of IVIVE methods for extrapolating results from in vitro assays to an in vivo equivalent dose that can be used in a quantitative risk assessment, there is lack of an internationally accepted guidance on IVIVE procedures and models.

4.e. Uncertainties.

80. The development of the nervous system, arguably the most complicated organ in the body, involves integration of intracellular, intercellular, interregional, and systemic interactions that occur in development-stage and regional specific manners. Due to the lack of knowledge of all possible molecular targets that, if disrupted, will alter nervous system development, the DNT IVB was designed to measure changes in some of the critical cellular processes downstream from potential molecular targets. Thus, the DNT IVB is different than in vitro batteries that assess one intracellular pathway (e.g., intracellular oestrogen signalling pathway). Thus, uncertainties in the evaluation of data from the DNT IVB (see previous sections) will be different compared to other batteries of in vitro assays. In addition to known uncertainties common to all in vitro assays (e.g., metabolism, untested chemical domains), the DNT IVB has additional uncertainties. These include the lack of assays for some critical developmental Kes (e.g., myelination), incomplete coverage of cell-to-cell interactions even in multicellular test systems (e.g., MEA, neurospheres), unknown neuronal subtypes, maturation stage, and complementary assays for only some processes (e.g., migration). The dearth of a consensus set of animal or human positive and negative reference chemicals adds uncertainty to the predictive nature of the DNT IVB and should be a high priority effort in the future (see Appendix A). These uncertainties are described in detail in previous sections. Importantly, these uncertainties must be judged in light of the problem formulation..

4.f. Usage in Hazard Assessments

81. It is important to note that this document does not provide guidance on follow-up testing or tiered testing that would inform use of in vitro data in hazard assessments. Such testing could include orthogonal assays (i.e., not part of the current DNT IVB) to confirm positives or negatives, or further chemical characterisation using alternative species (e.g., *C. elegans*, zebrafish), mechanistic testing in rodent, or use of the DNT or OECD TG 443 guidelines.

82. Development of a tiered decision framework, which has been previously advocated (Bal-Price et al. 2018; Bushnell et al. 2010; Coecke et al. 2007; Crofton et al. 2011; Fritsche 2017; Lein et al. 2007), has only recently been begun. One framework has been proposed (Masjosthusmann et al. 2020) that includes multiple tiers starting with toxicokinetics, then the DNT IVB, targeted follow-up in vitro testing, and targeted in vivo testing in rats only when necessary. This approach advocates using a WoE at each tier to determine if sufficient information is available for a hazard classification decision, prior to progression to the next tier. A draft tiered testing framework, which was developed to facilitate international discussion, is currently under review by the OECD.

83. Considering the currently limited use of data from the DNT IVB, it is recommended that regulatory jurisdictions create frameworks that are fit-for-purpose, taking into account scientific uncertainties and practical limitations of existing test methods, to incorporate the DNT IVB into their regulatory process that is reflective of their needs. This should include critical comparisons of the uncertainties and limitations of both the in vitro and in vivo test methods (e.g., NAFTA 2016; Paparella et al. 2020; Pham et al. 2020). Critical appraisals of in vivo based approaches may serve as an objective benchmark to be met or overcome with any new approach considered (NASEM, 2022).

84. Despite the uncertainties summarised above, data from the DNT IVB assays have been used to inform several hazard and risk decisions. For example, data from the NNF assay is being used, along with data from many other in vitro assays, to prioritise large number of perfluorinated chemicals for further testing (USEPA 2019). Prioritisation for flame retardant alternatives is being made based on the combination of data from both in vitro DNT assays and zebrafish (Behl et al. 2015). Data from multiple DNT IVB assays has been proposed for use in a WoE for organophosphates (USEPA 2020b). Data from the entire DNT IVB, were recently used by EFSA to develop IATA case studies for deltamethrin and flufenacet (EFSA et al. 2021). This effort resulted in an AOP-informed DNT risk assessment using all available hazard-related information (e.g., in vitro, toxicokinetics, epidemiology, in vivo

animal data) (see also Appendix D). A recent publication summarized a WoE analysis that integrated in vivo data on DL-glufosinate and data from from the DNT IVB assays for both DL- and L-glufosinate that utilized DNT IVB data in a WoE analysis that resulted in a waiver for a DNT guideline to waive the need for a new DNT guideline study on the enriched isomers (L-glufosinate ammonium, and L-glufosinate acid) (Dobreniecki et al., 2022). These examples illustrate how data from the DNT IVB can be applied to a wide variety of use contexts.

REFERENCES

Adhya D, Annuario E, Lancaster MA, Price J, Baron-Cohen S, Srivastava DP. 2018. Understanding the role of steroids in typical and atypical brain development: Advantages of using a "brain in a dish" approach. J Neuroendocrinol. 2018 Feb;30(2):e12547.

Algharably EAE, Di Consiglio E, Testai E, Kreutz R, Gundert-Remy U. 2021. Prediction of the dose range for adverse neurological effects of amiodarone in patients from an in vitro toxicity test by in vitro-in vivo extrapolation. Arch Toxicol. 2021 Apr;95(4):1433-1442

Arambula SE, McCarthy MM. Neuroendocrine-Immune Crosstalk Shapes Sex-Specific Brain Development. Endocrinology. 2020 Jun 1;161(6):bqaa055. doi: 10.1210/endocr/bqaa055.

Andersen ME, McMullen PD, Phillips MB, Yoon M, Pendse SN, Clewell HJ, Hartman JK, Moreau M, Becker RA, Clewell RA. 2019. Developing context appropriate toxicity testing approaches using new alternative methods (nams). Altex. 36(4):523-534.

Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, Mount DR, Nichols JW, Russom CL, Schmieder PK et al. 2010. Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. Environ Toxicol Chem. 29(3):730-741.

Aschner M, Ceccatelli S, Daneshian M, Fritsche E, Hasiwa N, Hartung T, Hogberg HT, Leist M, Li A, Mundi WR et al. 2017. Reference compounds for alternative test methods to indicate developmental neurotoxicity (dnt) potential of chemicals: Example lists and criteria for their selection and use. Altex. 34(1):49-74.

Bal-Price A, Crofton KM, Leist M, Allen S, Arand M, Buetler T, Delrue N, FitzGerald RE, Hartung T, Heinonen T et al. 2015. International stakeholder network (ISTNET): Creating a developmental neurotoxicity (DNT) testing road map for regulatory purposes. Arch Toxicol. 89(2):269-287.

Bal-Price A, Hogberg HT, Crofton KM, Daneshian M, FitzGerald RE, Fritsche E, Heinonen T, Hougaard Bennekou S, Klima S, Piersma AH et al. 2018. Recommendation on test readiness criteria for new approach methods in toxicology: Exemplified for developmental neurotoxicity. Altex. 35(3):306-352.

Ball K, Bouzom F, Scherrmann JM, Walther B, Decleves X. 2013. Physiologically based pharmacokinetic modelling of drug penetration across the blood-brain barrier--towards a mechanistic ivive-based approach. AAPS J. 15(4):913-932.

Barateiro A, Fernandes A. 2014. Temporal oligodendrocyte lineage progression: In vitro models of proliferation, differentiation and myelination. Biochim Biophys Acta. 1843(9):1917-1929.

Barbosa DJ, Capela JP, de Lourdes Bastos M, Carvalho F. 2015. In vitro models for neurotoxicology research. Toxicology Research. 4(4):801-842.

Barry C, Schmitz MT, Jiang P, Schwartz MP, Duffin BM, Swanson S, Bacher R, Bolin JM, Elwell AL, McIntosh BE et al. 2017. Species-specific developmental timing is maintained by pluripotent stem cells ex utero. Dev Biol. 423(2):101-110.

Beal MA, Gagne M, Kulkarni SA, Patlewicz G, Thomas RS, Barton-Maclaren TS. Implementing in vitro bioactivity data to modernize priority setting of chemical inventories. ALTEX. 2022;39(1):123-139.

Behl M, Hsieh JH, Shafer TJ, Mundy WR, Rice JR, Boyd WA, Freedman JH, Hunter ES, 3rd, Jarema KA, Padilla S et al. 2015. Use of alternative assays to identify and prioritize organophosphorus flame retardants for potential developmental and neurotoxicity. Neurotox. Teratol. 52(Pt B):181-193.

Behl M, Ryan K, Hsieh JH, Parham F, Shapiro AJ, Collins BJ, Sipes NS, Birnbaum LS, Bucher JR, Foster PMD et al. 2019. Screening for developmental neurotoxicity at the national toxicology program: The future is here. Toxicol Sci. 167(1):6-14.

Bernal J. Thyroid Hormones in Brain Development and Function. 2022. In: Feingold KR et al editors. Endotext [Internet]. South Dartmouth (MA): MDText.com, Inc.; 2000–. PMID: 25905404.

Bigbee JW, Sharma KV, Gupta JJ, Dupree JL. 1999. Morphogenic role for acetylcholinesterase in axonal outgrowth during neural development. Environmental health perspectives. 107 Suppl 1:81-87.

Borrell V, Gotz M. 2014. Role of radial glial cells in cerebral cortex folding. Curr Opin Neurobiol. 27:39-46.

Breen M, Ring CL, Kreutz A, Goldsmith MR, Wambaugh JF. 2021. High-throughput pbtk models for in vitro to in vivo extrapolation. Expert Opin Drug Metab Toxicol 17:903-921.

Brown JP, Hall D, Frank CL, Wallace K, Mundy WR, Shafer TJ. 2016. Evaluation of a microelectrode array-based assay for neural network ontogeny using training set chemicals. Toxicol Sci. 154(1):126-139.

Bushnell PJ, Kavlock RJ, Crofton KM, Weiss B, Rice DC. 2010. Behavioral toxicology in the 21st century: Challenges and opportunities for behavioral scientists. Summary of a symposium presented at the annual meeting of the neurobehavioral teratology society, june, 2009. Neurotoxicology and teratology. 32(3):313-328.

Cameron HA, Hazel TG, McKay RD. 1998. Regulation of neurogenesis by growth factors and neurotransmitters. J Neurobiol. 36(2):287-306.

Carmichael PL, Baltazar MT, Cable S, Cochrane S, Dent M, Li H, Middleton A, Muller I, Reynolds G, Westmoreland C, White A. Ready for regulatory use: NAMs and NGRA for chemical safety assurance. ALTEX. 2022;39(3):359–366.

Choi BH. 1986. Methylmercury poisoning of the developing nervous system: I. Pattern of neuronal migration in the cerebral cortex. Neurotoxicology. 7(2):591-600.

Coecke S, Goldberg AM, Allen S, Buzanska L, Calamandrei G, Crofton K, Hareng L, Hartung T, Knaut H, Honegger P et al. 2007. Workgroup report: Incorporating in vitro alternative methods for developmental neurotoxicity into international hazard and risk assessment strategies. Environmental health perspectives. 115(6):924-931.

Cote I, Andersen ME, Ankley GT, Barone S, Birnbaum LS, Boekelheide K, Bois FY, Burgoon LD, Chiu WA, Crawford-Brown D et al. 2016. The next generation of risk assessment multi-year study-highlights of findings, applications to risk assessment, and future directions. Environmental health perspectives. 124(11):1671-1682.

Cooper DJ, Zunino G, Bixby JL, Lemmon VP. 2017. Phenotypic screening with primary neurons to identify drug targets for regeneration and degeneration. Mol Cell Neurosci 80:161-169.

Cowan M, Petri WA, Jr. 2018. Microglia: Immune regulators of neurodevelopment. Front Immunol. 9:2576.

Crofton KM, Mundy WR. 2021. External scientific report on the interpretation of data from the developmental neurotoxicity in vitro testing assays for use in integrated approaches for testing and assessment. Efsa supporting publications, 18(10), 6924e. Https://doi.Org/10.2903/sp.Efsa.2021.En-6924.

Crofton KM. 2021. An estimate of the number of OECD 426, OECD 443 and EPA OPPTS 970.6300 guidance studies on developmental neurotoxicity (personal communication).

Crofton KM, Mundy WR, Lein PJ, Bal-Price A, Coecke S, Seiler AE, Knaut H, Buzanska L, Goldberg A. 2011. Developmental neurotoxicity testing: Recommendations for developing alternative methods for the screening and prioritization of chemicals. Altex. 28(1):9-15.

Crofton, K. M., A. Bassan, M. Behl, Y. G. Chushak, E. Fritsche, J. M. Gearhart, M. S. Marty, M. Mumtaz, M. Pavan, P. Ruiz, M. Sachana, R. Selvam, T. J. Shafer, L. Stavitskaya, D. T. Szabo, S. T. Szabo, R. R. Tice, D. Wilson, D. Woolley and G. J. Myatt (2022). "Current status and future directions for a neurotoxicity hazard assessment framework that integrates in silico approaches." Computational Toxicology 22: 100223.DOI: https://doi.org/10.1016/j.comtox.2022.100223

Croom EL, Shafer TJ, Evans MV, Mundy WR, Eklund CR, Johnstone AF, Mack CM, Pegram RA. 2015. Improving in vitro to in vivo extrapolation by incorporating toxicokinetic measurements: A case study of lindane-induced neurotoxicity. Toxicol Appl Pharmacol. 283(1):9-19.

Culbreth M, Nyffeler J, Willis C, Harrill JA. Optimization of Human Neural Progenitor Cells for an Imaging-Based High-Throughput Phenotypic Profiling Assay for Developmental Neurotoxicity Screening. Front Toxicol. 2022 Feb 16;3:803987.

Das KP, Barone S, Jr. 1999. Neuronal differentiation in pc12 cells is inhibited by chlorpyrifos and its metabolites: Is acetylcholinesterase inhibition the site of action? Toxicol Appl Pharmacol. 160(3):217-230.

Davis JA, Gift JS, Zhao QJ. 2011. Introduction to benchmark dose methods and u.S. Epa's benchmark dose software (bmds) version 2.1.1. Toxicol Appl Pharmacol. 254(2):181-191.

DeGroot DE, Swank A, Thomas RS, Strynar M, Lee MY, Carmichael PL, Simmons SO. 2018. Mrna transfection retrofits cell-based assays with xenobiotic metabolism. J Pharmacol Toxicol Methods. 92:77-94.

Delp J, Gutbier S, Klima S, Hoelting L, Pinto-Gil K, Hsieh JH, Aichem M, Klein K, Schreiber F, Tice RR et al. 2018. A high-throughput approach to identify specific neurotoxicants/ developmental toxicants in human neuronal cell function assays. Altex. 35(2):235-253.

Delsing L, Herland A, Falk A, Hicks R, Synnergren J, Zetterberg H. 2020. Models of the blood-brain barrier using ipsc-derived cells. Mol Cell Neurosci. 107:103533.

De Vries, GJ 2004. Minireview: Sex Differences in Adult and Developing Brains: Compensation, Compensation, Compensation, *Endocrinology*, 145:1063–1068,

Dobreniecki S, Mendez E, Lowit A, Freudenrich TM, Wallace K, Carpenter A, Wetmore BA, Kreutz A, Korol-Bexell E, Friedman KP, Shafer TJ. Integration of toxicodynamic and toxicokinetic new approach methods into a weight-of-evidence analysis for pesticide developmental neurotoxicity assessment: A case-study with DL- and L-glufosinate. Regul Toxicol Pharmacol. 2022 Apr 9;131:105167. doi: 10.1016/j.yrtph.2022.105167. Epub ahead of print. PMID: 35413399.

EFSA, Hernández-Jerez A, Adriaanse P, Aldrich A, Berny P, Coja T, Duquesne S, Focks A, Marinovich M, Millet M et al. 2021. Development of integrated approaches to testing and assessment (iata) case studies on developmental neurotoxicity (dnt) risk assessment. EFSA Journal. 19(6):e06599.

Escher BI, Henneberger L, König M, Schlichting R, Fischer FC. Cytotoxicity Burst? Differentiating Specific from Nonspecific Effects in Tox21 in Vitro Reporter Gene Assays. Environ Health Perspect. 2020 Jul;128(7):77007. doi: 10.1289/EHP6664. Epub 2020 Jul 23. PMID: 32700975; PMCID: PMC7377237

Ferguson CS, Tyndale RF. 2011. Cytochrome p450 enzymes in the brain: Emerging evidence of biological significance. Trends Pharmacol Sci. 32(12):708-714.

Filer DL, Kothiya P, Setzer RW, Judson RS, Martin MT. 2017. Tcpl: The toxcast pipeline for high-throughput screening data. Bioinformatics. 33(4):618-620.

Frank CL, Brown JP, Wallace K, Mundy WR, Shafer TJ. 2017. From the cover: Developmental neurotoxicants disrupt activity in cortical networks on microelectrode arrays: Results of screening 86 compounds during neural network formation. Toxicol Sci. 160(1):121-135.

Fritsche E. 2017. Report on integrated testing strategies for the identification and evaluation of chemical hazards associated with the developmental neurotoxicity (dnt), to facilitate discussions at the joint efsa. In OECD Workshop on DNT. Env/jm/mono (2017) (vol. 4).

Fritsche E, Grandjean P, Crofton KM, Aschner M, Goldberg A, Heinonen T, et al. 2018. Consensus statement on the need for innovation, transition and implementation of developmental neurotoxicity (DNT) testing for regulatory purposes. Toxicol Appl Pharmacol 354:3-6.

Fritsche E, Alm H, Baumann J, Geerts L, Håkansson H, Masjosthusmann S, Witters H. 2015. Literature review on in vitro and alternative developmental neurotoxicity (dnt) testing methods. EFSA Supporting Publications. 12(4):778E.

Fritsche E, Crofton KM, Hernandez AF, Hougaard Bennekou S, Leist M, Bal-Price A, Reaves E, Wilks MF, Terron A, Solecki R et al. 2017. Oecd/efsa workshop on developmental neurotoxicity (dnt): The use of non-animal test methods for regulatory purposes. Altex. 34(2):311-315.

Gingrich J, Filipovic D, Conolly R, Bhattacharya S, Veiga-Lopez A. Pregnancy-specific physiologically-based toxicokinetic models for bisphenol A and bisphenol S. Environ Int. 2021 Feb;147:106301.

Goodman LS. 1996. Goodman and gilman's the pharmacological basis of therapeutics. McGraw-Hill New York.

Grandjean P, Landrigan PJ. 2006. Developmental neurotoxicity of industrial chemicals. Lancet. 368(9553):2167-2178.

Grandjean P, Landrigan PJ. 2014. Neurobehavioural effects of developmental toxicity. Lancet Neurol. 13(3):330-338.

Griesinger C, Desprez B, Coecke S, Casey W, Zuang V. 2016. Validation of alternative in vitro methods to animal testing: Concepts, challenges, processes and tools. Adv Exp Med Biol. 856:65-132.

Guo BQ, Yan CH, Cai SZ, Yuan XB, Shen XM. 2013. Low level prenatal exposure to methylmercury disrupts neuronal migration in the developing rat cerebral cortex. Toxicology. 304:57-68.

Hardy A, Benford D, Halldorsson T, Jeger MJ, Knutsen KH, More S, Mortensen A, Naegeli H, Noteborn H, Ockleford C et al. 2017. Update: Use of the benchmark dose approach in risk assessment. EFSA J. 15(1):e04658.

Harrill JA, Freudenrich T, Wallace K, Ball K, Shafer TJ, Mundy WR. 2018. Testing for developmental neurotoxicity using a battery of in vitro assays for key cellular events in neurodevelopment. Toxicol Appl Pharmacol. 354:24-39.

Harrill JA, Freudenrich TM, Machacek DW, Stice SL, Mundy WR. 2010. Quantitative assessment of neurite outgrowth in human embryonic stem cell-derived hn2 cells using automated high-content image analysis. Neurotoxicology. 31(3):277-290.

Harrill JA, Freudenrich TM, Robinette BL, Mundy WR. 2011a. Comparative sensitivity of human and rat neural cultures to chemical-induced inhibition of neurite outgrowth. Toxicol Appl Pharmacol. 256(3):268-280.

Harrill JA, Robinette BL, Mundy WR. 2011b. Use of high content image analysis to detect chemical-induced changes in synaptogenesis in vitro. Toxicol In Vitro. 25(1):368-387.

Harry GJ, Billingsley M, Bruinink A, Campbell IL, Classen W, Dorman DC, Galli C, Ray D, Smith RA, Tilson HA. 1998. In vitro techniques for the assessment of neurotoxicity. Environmental health perspectives. 106 Suppl 1:131-158.

Harry GJ, Tiffany-Castiglioni E. 2005. Evaluation of neurotoxic potential by use of in vitro systems. Expert Opin Drug Metab Toxicol. 1(4):701-713.

Hedlund E, Gustafsson JA, Warner M. 2001. Cytochrome p450 in the brain; a review. Curr Drug Metab. 2(3):245-263.

Hornung MW, Korte JJ, Olker JH, Denny JS, Knutsen C, Hartig PC, Cardon MC, Degitz SJ. 2018. Screening the toxcast phase 1 chemical library for inhibition of deiodinase type 1 activity. Toxicol Sci. 162(2):570-581.

Howard AS, Bucelli R, Jett DA, Bruun D, Yang D, Lein PJ. 2005. Chlorpyrifos exerts opposing effects on axonal and dendritic growth in primary neuronal cultures. Toxicol Appl Pharmacol. 207(2):112-124.

1311 Hsieh JH, Sedykh A, Huang R, Xia M, Tice RR. 2015. A data analysis pipeline accounting for artifacts in tox21 quantitative
1312 high-throughput screening assays. J Biomol Screen. 20(7):887-897.
1313 Jones L, Fischer I, Levitt P. 1996. Nonuniform alteration of dendritic development in the cerebral cortex following prenatal
1314 cocaine exposure. Cereb Cortex. 6(3):431-445.
1315 Judson R, Houck K, Martin M, Richard AM, Knudsen TB, Shah I, Little S, Wambaugh J, Setzer RW, Kothiya P et al. 2016.
1316 Analysis of the effects of cell stress and cytotoxicity on in vitro assay activity across a diverse chemical and assay space.
1317 Toxicol Sci. 153(2):409.
1318 Judson RS, Magpantay FM, Chickarmane V, Haskell C, Tania N, Taylor J, Xia M, Huang R, Rotroff DM, Filer DL, Houck KA,
1319 Martin MT, Sipes N, Richard AM, Mansouri K, Setzer RW, Knudsen TB, Crofton KM, Thomas RS. Integrated Model of
1320 Chemical Perturbations of a Biological Pathway Using 18 In Vitro High-Throughput Screening Assays for the Estrogen
1321 Receptor. Toxicol Sci. 2015 Nov;148(1):137-54.
1322 Judson R, Kavlock R, Martin M, Reif D, Houck K, Knudsen T, Richard A, Tice RR, Whelan M, Xia M. 2013. Perspectives on
1323 validation of high-throughput assays supporting 21st century toxicity testing. Altex. 30(1):51.
1324 Kartvelishvily E, Shleper M, Balan L, Dumin E, Wolosker H. 2006. Neuron-derived d-serine release provides a novel means to
1325 activate n-methyl-d-aspartate receptors. J Biol Chem. 281(20):14151-14162.
1326 Kavlock RJ, Bahadori T, Barton-Maclaren TS, Gwinn MR, Rasenberg M, Thomas RS. 2018. Accelerating the pace of chemical
1327 risk assessment. Chem Res Toxicol. 31(5):287-290.
1328 Kirkland D, Aardema M, Henderson L, Muller L. 2005. Evaluation of the ability of a battery of three in vitro genotoxicity tests
1329 to discriminate rodent carcinogens and non-carcinogens i. Sensitivity, specificity and relative predictivity. Mutat Res.
1330 584(1-2):1-256.
1331 Klocke C, Lein PJ. 2020. Evidence implicating non-dioxin-like congeners as the key mediators of polychlorinated biphenyl
1332 (PCB) developmental neurotoxicity. Int J Mol Sci. 21(3):1013. doi: 10.3390/ijms21031013.
1333 Klose J, Tigges J, Masjosthusmann S, Schmuck K, Bendt F, Hübenthal U, Petzsch P, Köhrer K, Koch K, Fritsche E. TBBPA
1334 targets converging key events of human oligodendrocyte development resulting in two novel AOPs. ALTEX.
1335 2021;38(2):215-234.
1336 Kramer NI, Di Consiglio E, Blaauboer BJ, Testai E. 2015. Biokinetics in repeated-dosing in vitro drug toxicity studies. Toxicol
1337 In Vitro. 30(1 Pt A):217-224.
1338 Krebs A, Waldmann T, Wilks MF, van Vugt-Lussenburg BMA, Van der Burg B, Terron A, Steger-Hartmann T, Ruegg J, Rovida
1339 C, Pedersen E. 2019. Template for the description of cell-based toxicological test methods to allow evaluation and
1340 regulatory use of the data. . Altex. 36(4):682-699. [correction appears in altex. 2020;37(1):164]
1341 Krug AK, Balmer NV, Matt F, Schonenberger F, Merhof D, Leist M. 2013. Evaluation of a human neurite growth assay as
1342 specific screen for developmental neurotoxicants. Arch Toxicol. 87(12):2215-2231.
1343 Kunimoto M, Suzuki T. 1997. Migration of granule neurons in cerebellar organotypic cultures is impaired by methylmercury.
1344 Neurosci Lett. 226(3):183-186.
1345 Lein P, Locke P, Goldberg A. 2007. Meeting report: Alternatives for developmental neurotoxicity testing. Environmental
1346 health perspectives. 115(5):764-768.
1347 Lein P, Silbergeld E, Locke P, Goldberg AM. 2005. In vitro and other alternative approaches to developmental neurotoxicity
1348 testing (dnt). Environmental toxicology and pharmacology. 19(3):735-744.
1349 Leys C, Ley C, Klein O, Bernard P, Licata L. 2013. Detecting outliers: Do not use standard deviation around the mean, use
1350 absolute deviation around the median. Journal of experimental social psychology. 49(4):764-766.
1351 Loser D, Schaefer J, Danker T, Möller C, Brüll M, Suciu I, Ückert AK, Klima S, Leist M, Kraushaar U. Human neuronal signaling
1352 and communication assays to assess functional neurotoxicity. Arch Toxicol. 2021 95(1):229-252.
1353 Makris SL, Raffaele K, Allen S, Bowers WJ, Hass U, Alleva E, Calamandrei G, Sheets L, Amcoff P, Delrue N et al. 2009. A
1354 retrospective performance assessment of the developmental neurotoxicity study in support of oecd test guideline 426.
1355 Environmental health perspectives. 117(1):17-25.
1356 Malo N, Hanley JA, Cerquozzi S, Pelletier J, Nadon R. 2006. Statistical practice in high-throughput screening data analysis.
1357 Nat Biotechnol. 24(2):167-175.
1358 Martin MM, Baker NC, Boyes WK, Carstens KE, Culbreth ME, Gilbert ME, et al. 2021. An expert driven literature review of
1359 “negative” reference chemicals for developmental neurotoxicity (DNT) assay evaluation.(
1360 https://epa.Figshare.Com/articles/poster/an_expert-
1361 driven_literature_review_of_negative_reference_chemicals_for_developmental_neurotoxicity_dnt_assay_evaluation/
1362 14226326

Masjosthusmann S, Becker D, Petzuch B, Klose J, Siebert C, Deenen R, Barenys M, Baumann J, Dach K, Tigges J et al. 2018. A transcriptome comparison of time-matched developing human, mouse and rat neural progenitor cells reveals human uniqueness. Toxicol Appl Pharmacol. 354:40-55.

Masjosthusmann S, Blum J, Bartmann K, Dolde X, Holzer AK, Stürzl LC, Keßel EH, Förster N, Dönmez A, Klose J. 2020. Establishment of an a priori protocol for the implementation and interpretation of an in-vitro testing battery for the assessment of developmental neurotoxicity. EFSA Supporting Publications. 17(10):1938E.

Mayor R, Theveneau E. 2013. The neural crest. Development. 140(11):2247-2251.

Mileson BE, Ferenc SA. 2001. Methods to identify and characterize developmental neurotoxicity for human health risk assessment: Overview. Environmental health perspectives. 109 Suppl 1:77-78.

Miller MD, Crofton KM, Rice DC, Zoeller RT. 2009. Thyroid-disrupting chemicals: Interpreting upstream biomarkers of adverse outcomes. Environmental health perspectives. 117(7):1033-1041.

Miller MW. 1986. Effects of alcohol on the generation and migration of cerebral cortical neurons. Science. 233(4770):1308-1311.

Minoux M, Rijli FM. 2010. Molecular mechanisms of cranial neural crest cell migration and patterning in craniofacial development. Development. 137(16):2605-2621.

Moors M, Rockel TD, Abel J, Cline JE, Gassmann K, Schreiber T, Schuwald J, Weinmann N, Fritsche E. 2009. Human neurospheres as three-dimensional cellular systems for developmental neurotoxicity testing. Environmental health perspectives. 117(7):1131-1138.

Mundy WR, Freudenrich TM, Crofton KM, DeVito MJ. 2004. Accumulation of pbde-47 in primary cultures of rat neocortical cells. Toxicol Sci. 82(1):164-169.

Mundy WR, Padilla S, Breier JM, Crofton KM, Gilbert ME, Herr DW, Jensen KF, Radio NM, Raffaele KC, Schumacher K et al. 2015. Expanding the test set: Chemicals with potential to disrupt mammalian brain development. Neurotoxicology and teratology. 52(Pt A):25-35.

Mundy WR, Radio NM, Freudenrich TM. 2010. Neuronal models for evaluation of proliferation in vitro using high content screening. Toxicology. 270(2-3):121-130.

NAFTA (2016) Developmental Neurotoxicity Study Guidance Document, Technical Working Group on Pesticides (TWG), North American Free Trade Agreement (NAFTA), December, 2016. p.133. https://www.epa.gov/pesticide-science-and-assessing-pesticide-risks/developmental-neurotoxicity-study-guidance

NASEM (2022) Variability and Relevance of Current Laboratory Mammalian Toxicity Tests and Expectations for New Approach Methods (NAMs) for use in Human Health Risk Assessment. Ad HOC Committee, National Academy of Sciences, Engineering and Medicine. https://www.nationalacademies.org/our-work/variability-and-relevance-of-current-laboratory-mammalian-toxicity-tests-and-expectations-for-new-approach-methods--nams--for-use-in-human-health-risk-assessment

Nimtz L, Hartmann J, Tigges J, Masjosthusmann S, Schmuck M, Kessel E, Theiss S, Kohrer K, Petzsch P, Adjaye J et al. 2020. Characterization and application of electrically active neuronal networks established from human induced pluripotent stem cell-derived neural progenitor cells for neurotoxicity evaluation. Stem Cell Res. 45:101761.

Nimtz L, Klose J, Masjosthusmann S, Barenys M, Fritsche E. 2019. The neurosphere assay as an in vitro method for developmental neurotoxicity (dnt) evaluation. In cell culture techniques (2nd edtion) . Aschner,m. & costa, l. (eds), humana, new york, ny. Pp. 141-168.

Nyffeler J, Dolde X, Krebs A, Pinto-Gil K, Pastor M, Behl M, Waldmann T, Leist M. 2017a. Combination of multiple neural crest migration assays to identify environmental toxicants from a proof-of-concept chemical library. Arch Toxicol. 91(11):3613-3632.

Nyffeler J, Karreman C, Leisner H, Kim YJ, Lee G, Waldmann T, Leist M. 2017b. Design of a high-throughput human neural crest cell migration assay to indicate potential developmental toxicants. Altex. 34(1):75-94.

Odawara A, Katoh H, Matsuda N, Suzuki I. 2016. Physiological maturation and drug responses of human induced pluripotent stem cell-derived cortical neuronal networks in long-term culture. Sci Rep. 6:26181.

OECD. 2007. Test no. 426: Developmental neurotoxicity study. OECD Publishing.

OECD. 2008. Oecd series on testing and assessment, number 89: Retrospective performance assessment of the test guideline 426 on developmental neurotoxicity. OECD Publishing.

OECD. 2016. Guidance document for the use of adverse outcome pathways in developing integrated approaches to testing and assessment (iata). .

OECD. 2017a. Annex 1. Report of the oecd/efsa workshop on developmental neurotoxicity (dnt): The use of non-animal test methods for regualtory purposes. Background document on integrated testing strategies for the identification and evaluation of chemical hazards associated with the developmental neurotoxicity (dnt), to facilitate discussions at the joint efsa/oecd workshop on dnt. Series on testing and assessment. No. 26. Env/jm/mono(2017)4/ann1.

OECD. 2017b. Guidance document for describing non-guideline in vitro test methods. Series on testing and assessment no. 211. Oecd publishing, paris. .

OECD. 2018. Guidance document on good in vitro method practices (givimp).

OECD. 2019. Guiding principles and key elements for establishing a weight of evidence for chemical assessment. Series on testing and assessment no. 311, environment, health and safety division, environment directorate. Oecd publishing, paris. .

Ogunshola OO, Antic A, Donoghue MJ, Fan SY, Kim H, Stewart WB, Madri JA, Ment LR. 2002. Paracrine and autocrine functions of neuronal vascular endothelial growth factor (vegf) in the central nervous system. J Biol Chem. 277(13):11410-11415.

Ojeda J, Avila A. 2019. Early actions of neurotransmitters during cortex development and maturation of reprogrammed neurons. Front Synaptic Neurosci. 11:33. doi: 10.3389/fnsyn.2019.00033

Olker JH, Korte JJ, Denny JS, Hartig PC, Cardon MC, Knutsen CN, Kent PM, Christensen JP, Degitz SJ, Hornung MW. 2019. Screening the toxcast phase 1, phase 2, and e1k chemical libraries for inhibitors of iodothyronine deiodinases. Toxicol Sci. 168(2):430-442.

Ortega MC, Bribian A, Peregrin S, Gil MT, Marin O, de Castro F. 2012. Neuregulin-1/erbb4 signaling controls the migration of oligodendrocyte precursor cells during development. Exp Neurol. 235(2):610-620.

Paini A, Leonard JA, Joossens E, Bessems JGM, Desalegn A, Dorne JL, Gosling JP, Heringa MB, Klaric M, Kliment T et al. 2019. Next generation physiologically based kinetic (ng-pbk) models in support of regulatory decision making. Comput Toxicol. 9:61-72.

Paparella M, Bennekou SH, Bal-Price A. 2020. An analysis of the limitations and uncertainties of in vivo developmental neurotoxicity testing and assessment to identify the potential for alternative approaches. Reprod Toxicol. 96:327-336.

Parish ST, Aschner M, Casey W, Corvaro M, Embry MR, Fitzpatrick S, Kidd D, Kleinstreuer NC, Lima BS, Settivari RS, Wolf DC, Yamazaki D, Boobis A. An evaluation framework for new approach methodologies (NAMs) for human health safety assessment. Regul Toxicol Pharmacol. 2020 Apr;112:104592.

Patterson EA, Whelan MP, Worth AP. The role of validation in establishing the scientific credibility of predictive toxicology approaches intended for regulatory application. Comput Toxicol. 2021 Feb;17:100144.

Paul-Friedman K, Martin M, Crofton KM, Hsu CW, Sakamuru S, Zhao J, Xia M, Huang R, Stavreva DA, Soni V et al. 2019. Limited chemical structural diversity found to modulate thyroid hormone receptor in the tox21 chemical library. Environmental health perspectives. 127(9):97009.

Paul Friedman K, Watt ED, Hornung MW, Hedge JM, Judson RS, Crofton KM, Houck KA, Simmons SO. 2016. Tiered high-throughput screening approach to identify thyroperoxidase inhibitors within the toxcast phase i and ii chemical libraries. Toxicol Sci. 151(1):160-180.

Paul Friedman K. 2022. Personnal communication, February 2022.

Ly Pham L, Watford S, Pradeep P, Martin MT, Thomas R, Judson R, Setzer RW, Paul Friedman K. Variability in in vivo studies: Defining the upper limit of performance for predictions of systemic effect levels. Comput Toxicol. 1;15:1-100126.

Prince LM, Aschner M, Bowman AB. 2019. Human-induced pluripotent stems cells as a model to dissect the selective neurotoxicity of methylmercury. Biochim Biophys Acta Gen Subj. 1863(12):129300.

Proenca S, Escher BI, Fischer FC, Fisher C, Gregoire S, Hewitt NJ, Nicol B, Paini A, Kramer NI. 2021. Effective exposure of chemicals in in vitro cell systems: A review of chemical distribution models. Toxicol In Vitro. 73:105133.

Qiao D, Seidler FJ, Tate CA, Cousins MM, Slotkin TA. 2003. Fetal chlorpyrifos exposure: Adverse effects on brain cell development and cholinergic biomarkers emerge postnatally and continue into adolescence and adulthood. Environmental health perspectives. 111(4):536-544.

Radio NM, Mundy WR. 2008. Developmental neurotoxicity testing in vitro: Models for assessing chemical effects on neurite outgrowth. Neurotoxicology. 29(3):361-376.

Rees DC, Francis EZ, Kimmel CA. 1990. Scientific and regulatory issues relevant to assessing risk for developmental neurotoxicity: An overview. Neurotoxicology and teratology. 12(3):175-181.

Reif DM, Martin MT, Tan SW, Houck KA, Judson RS, Richard AM, Knudsen TB, Dix DJ, Kavlock RJ. 2010. Endocrine profiling and prioritization of environmental chemicals using toxcast data. Environmental health perspectives. 118(12):1714-1720.

Rice D, Barone S, Jr. 2000. Critical periods of vulnerability for the developing nervous system: Evidence from humans and animal models. Environmental health perspectives. 108 Suppl 3:511-533.

Richard AM, Huang R, Waidyanatha S, Shinn P, Collins BJ, Thillainadarajah I, Grulke CM, Williams AJ, Lougee RR, Judson RS et al. 2020. The tox21 10k compound library: Collaborative chemistry advancing toxicology. Chem Res Toxicol. 34(2):189-216

Richard AM, Judson RS, Houck KA, Grulke CM, Volarath P, Thillainadarajah I, Yang C, Rathman J, Martin MT, Wambaugh JF et al. 2016. Toxcast chemical landscape: Paving the road to 21st century toxicology. Chem Res Toxicol. 29(8):1225-1251.

Robbins RD, Prasain N, Maier BF, Yoder MC, Mirmira RG. 2010. Inducible pluripotent stem cells: Not quite ready for prime time? Curr Opin Organ Transplant. 15(1):61-67.

Rodier PM. 1994. Vulnerable periods and processes during central nervous system development. Environmental health perspectives. 102 Suppl 2:121-124.

Ryan KR, Sirenko O, Parham F, Hsieh JH, Cromwell EF, Tice RR, Behl M. 2016 Neurite outgrowth in human induced pluripotent stem cell-derived neurons as a high-throughput screen for developmental neurotoxicity or neurotoxicity. Neurotoxicology. 53:271-281.

Saavedra L, Wallace K, Freudenrich T, Mall M, Mundy W, Davila J, Shafer T, Wernig M, Haag D. 2021. Comparison of acute effects of neurotoxic compounds on network activity in human and rodent neural cultures. Toxicol Sci. 180:295-312.

Sachana M, Willett C, Pistollato F, Bal-Price A. 2021. The potential of mechanistic information organised within the aop framework to increase regulatory uptake of the developmental neurotoxicity (dnt) in vitro battery of assays. Reprod Toxicol 103:159-170.

Sachana M, Bal-Price A, Crofton KM, Bennekou SH, Shafer TJ, Behl M, Terron A. 2019. International regulatory and scientific effort for improved developmental neurotoxicity testing. Toxicol Sci. 167(1):45-57.

Sachana M, Leinala E. 2017. Approaching chemical safety assessment through application of integrated approaches to testing and assessment: Combining mechanistic information derived from adverse outcome pathways and alternative methods. Applied In Vitro Toxicology. 3(3):227-233.

Sakuratani Y, Horie M, Leinala E. 2018. Integrated approaches to testing and assessment: Oecd activities on the development and use of adverse outcome pathways and case studies. Basic & Clinical Pharmacology & Toxicology. 123:20-28.

Sanes DH, Reh TA, Harris WA. 2011. Development of the nervous system. Academic Press.

Sass JB, Haselow DT, Silbergeld EK. 2001. Methylmercury-induced decrement in neuronal migration may involve cytokine-dependent mechanisms: A novel method to assess neuronal movement in vitro. Toxicol Sci. 63(1):74-81.

Schmidt BZ, Lehmann M, Gutbier S, Nembo E, Noel S, Smirnova L, Forsby A, Hescheler J, Avci HX, Hartung T et al. 2017. In vitro acute and developmental neurotoxicity screening: An overview of cellular platforms and high-throughput technical possibilities. Arch Toxicol. 91(1):1-33.

Schreiber T, Gassmann K, Gotz C, Hubenthal U, Moors M, Krause G, Merk HF, Nguyen NH, Scanlan TS, Abel J et al. 2010. Polybrominated diphenyl ethers induce developmental neurotoxicity in a human in vitro model: Evidence for endocrine disruption. Environmental health perspectives. 118(4):572-578.

Shafer TJ, Brown JP, Lynch B, Davila-Montero S, Wallace K, Friedman KP. 2019. Evaluation of chemical effects on network formation in cortical neurons grown on microelectrode arrays. Toxicol Sci. 169(2):436-455.

Shafer TJ, Hughes MF. 2010. Accumulation of pyrethroid compounds in primary cultures from rat cortex. Toxicol In Vitro. 24(7):2053-2057.

Sild M, Ruthazer ES. 2011. Radial glia: Progenitor, pathway, and partner. Neuroscientist. 17(3):288-302.

Sipes NS, Wambaugh JF, Pearce R, Auerbach SS, Wetmore BA, Hsieh JH, et al. 2017. An intuitive approach for predicting potential human health risk with the tox21 10k library. Environ Sci Technol 51:10786-10796.

Smirnova L, Hogberg HT, Leist M, Hartung T. 2014. Developmental neurotoxicity - challenges in the 21st century and in vitro opportunities. Altex. 31(2):129-156.

Spinu, N., Bal-Price, A., Cronin, M.T.D. et al. 2019. Development and analysis of an adverse outcome pathway network for human neurotoxicity. Arch Toxicol 93:2759–2772.

Stiegler NV, Krug AK, Matt F, Leist M. 2011. Assessment of chemical-induced impairment of human neurite outgrowth by multiparametric live cell imaging in high-density cultures. Toxicol Sci. 121(1):73-87.

Strickland JD, Martin MT, Richard AM, Houck KA, Shafer TJ. 2018. Screening the toxcast phase ii libraries for alterations in network function using cortical neurons grown on multi-well microelectrode array (mwmea) plates. Arch Toxicol. 92(1):487-500.

Thomas RS, Paules RS, Simeonov A, Fitzpatrick SC, Crofton KM, Casey WM, Mendrick DL. 2018. The us federal tox21 program: A strategic and operational plan for continued leadership. Altex. 35(2):163-168.

Thomas RS, Philbert MA, Auerbach SS, Wetmore BA, Devito MJ, Cote I, Rowlands JC, Whelan MP, Hays SM, Andersen ME et al. 2013. Incorporating new technologies into toxicity testing and risk assessment: Moving from 21st century vision to a data-driven framework. Toxicol Sci. 136(1):4-18.

Thorne N, Auld DS, Inglese J. Apparent activity in high-throughput screening: origins of compound-dependent assay interference. Curr Opin Chem Biol. 2010 (3):315-24.

Tice RR, Austin CP, Kavlock RJ, Bucher JR. 2013. Improving the human hazard characterization of chemicals: A tox21 update. Environmental health perspectives. 121(7):756-765.

Tingling JD, Bake S, Holgate R, Rawlings J, Nagsuk PP, Chandrasekharan J, Schneider SL, Miranda RC. 2013. Cd24 expression identifies teratogen-sensitive fetal neural stem cell subpopulations: Evidence from developmental ethanol exposure and orthotopic cell transfer models. PLoS One. 8(7):e69560.

Tohyama C. 2016. Developmental neurotoxicity test guidelines: Problems and perspectives. J Toxicol Sci. 41(Spec):SP69-SP79.

Tollefsen KE, Scholz S, Cronin MT, Edwards SW, de Knecht J, Crofton K, Garcia-Reyero N, Hartung T, Worth A, Patlewicz G. 2014. Applying adverse outcome pathways (aops) to support integrated approaches to testing and assessment (iata). Regul Toxicol Pharmacol. 70(3):629-640.

Tsuji R, Crofton KM. 2012. Developmental neurotoxicity guideline study: Issues with methodology, evaluation and regulation. Congenit Anom (Kyoto). 52(3):122-128.

Tucker EK, Nowak RA. 2018. Methylmercury alters proliferation, migration, and antioxidant capacity in human htr8/sv-neo trophoblast cells. Reprod Toxicol. 78:60-68.

USEPA. 1998. Health effects guidelines oppts 970.6300: Developmental neurotoxicity study.

USEPA. 2012. Scientific issues associated with chlorpyrifos health effects. Fifra scientific advisory panel open meeting, arlington, va, april 10 – 13, 2012. .

USEPA. 2019. EPA's per- and polyfluoroalkyl substances (pfas) action plan. Epa 823r18004. February 2019.

USEPA. 2020a. Chemical safety for sustainability - strategic research action plan. Us epa office of research and development. Epa 601 k20001 march 2020. Www.Epa.Gov/research.

USEPA. 2020b. The use of new approach methodologies (nams) to derive extrapolation factors and evaluate developmental neurotoxicity for human health risk assessment. Fifra scientific advisory panel. September 15-17, 2020. .

Valdivia P, Martin M, LeFew WR, Ross J, Houck KA, Shafer TJ. 2014. Multi-well microelectrode array recordings detect neuroactivity of toxcast compounds. Neurotoxicology. 44:204-217.

Vorrink SU, Zhou Y, Ingelman-Sundberg M, Lauschke VM. 2018. Prediction of drug-induced hepatotoxicity using long-term stable primary hepatic 3d spheroid cultures in chemically defined conditions. Toxicol Sci. 163(2):655-665.

Wambaugh JF, Wetmore BA, Ring CL, Nicolas CI, Pearce RG, Honda GS, Dinallo R, Angus D, Gilbert J, Sierra T et al. 2019. Assessing toxicokinetic uncertainty and variability in risk prioritization. Toxicol Sci. 172(2):235-251.

Wang J, Hallinger DR, Murr AS, Buckalew AR, Simmons SO, Laws SC, Stoker TE. 2018. High-throughput screening and quantitative chemical ranking for sodium-iodide symporter inhibitors in toxcast phase i chemical library. Environ Sci Technol. 52(9):5417-5426.

Watt ED, Judson RS. 2018. Uncertainty quantification in toxcast high throughput screening. PLoS One. 13(7):e0196963.

Weinberg AM. 1972. Science and trans-science. Science. 177(4045):211.

Wetmore BA, Wambaugh JF, Ferguson SS, Sochaski MA, Rotroff DM, Freeman K, Clewell HJ, 3rd, Dix DJ, Andersen ME, Houck KA et al. 2012. Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. Toxicol Sci. 125(1):157-174.

Wong MK, Li EW, Adam M, Selvaganapathy PR, Raha S. Establishment of an in vitro placental barrier model cultured under physiologically relevant oxygen levels. Mol Hum Reprod. 2020 May 15;26(5):353-365. Yang D, Kim KH, Phimister A, Bachstetter AD, Ward TR, Stackman RW, Mervis RF, Wisniewski AB, Klein SL, Kodavanti PR et al. 2009. Developmental exposure to polychlorinated biphenyls interferes with experience-dependent dendritic plasticity and ryanodine receptor expression in weanling rats. Environmental health perspectives. 117(3):426-435.

Yang D, Lauridsen H, Buels K, Chi LH, La Du J, Bruun DA, Olson JR, Tanguay RL, Lein PJ. 2011. Chlorpyrifos-oxon disrupts zebrafish axonal growth and motor behavior. Toxicol Sci. 121(1):146-159.

Yu S, Patchev AV, Wu Y, Lu J, Holsboer F, Zhang JZ, Sousa N, Almeida OF. 2010. Depletion of the neural precursor cell pool by glucocorticoids. Ann Neurol. 67(1):21-30.

van der Zalm AJ, Barroso J, Browne P, Casey W, Gordon J, Henry TR, Kleinstreuer NC, Lowit AB, Perron M, Clippinger AJ. A framework for establishing scientific confidence in new approach methodologies. Arch Toxicol. 2022 Aug 20. doi: 10.1007/s00204-022-03365-4. Epub ahead of print.

Zhang JH, Chung TD, Oldenburg KR. 1999. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. J Biomol Screen. 4(2):67-73.

Zimmer B, Lee G, Balmer NV, Meganathan K, Sachinidis A, Studer L, Leist M. 2012. Evaluation of developmental toxicants and signaling pathways in a functional test based on the migration of human neural crest cells. Environmental health perspectives. 120(8):1116-1122.

Zoeller RT, Rovet J. 2004. Timing of thyroid hormone action in the developing brain: Clinical observations and experimental findings. J Neuroendocrinol. 16(10):809-818.

1582

List of Appendices and Filenames

Appendix	Title	Filenames
A	Proposed Positive and Negative Control Compounds for Use in Performance Evaluations of DNT IVB Assays	Appendix A
	• List of proposed chemicals	Appendix A.1A
	• Background information	Appendix A.1B
B	Assay Names	Appendix B Number
	• NPC1	Appendix B.1
	• NPC2a	Appendix B.2
	• NPC2b	Appendix B.2
	• NPC2c	Appendix B.2
	• NPC3	Appendix B.2
	• NPC4	Appendix B.2
	• NPC5	Appendix B.2
	• UKN2	Appendix B.3
	• UKN4	Appendix B.4
	• UKN5	Appendix B.5
	• Cortical Initiation	Appendix B.6
	• Cortical Maturation	Appendix B.7
	• Cortical Synapto	Appendix B.7
	• Cortical MEA	Appendix B.8
	• hN initiation	Appendix B.9
	• hNP1 Apop	Appendix B.10
	• hNP1 Prolif	Appendix B.10
C	Adverse Outcome Pathway as an Underlying Framework for Developing In Vitro DNT Testing Strategies	Appendix C
D	List of DNT IATAs	Appendix D
E	List of all chemicals tested in the DNT IVB assays	Appendix E

1583